

University of Groningen

Balanced Realization and Model Order Reduction for Nonlinear Systems based on Singular Value Analysis

Fujimoto, Kenji; Scherpen, Jacqueliën M. A.

Published in:
SIAM Journal on Control and Optimization

DOI:
[10.1137/070695332](https://doi.org/10.1137/070695332)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2010

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Fujimoto, K., & Scherpen, J. M. A. (2010). Balanced Realization and Model Order Reduction for Nonlinear Systems based on Singular Value Analysis. *SIAM Journal on Control and Optimization*, 48(7), 4591-4623. <https://doi.org/10.1137/070695332>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

BALANCED REALIZATION AND MODEL ORDER REDUCTION FOR NONLINEAR SYSTEMS BASED ON SINGULAR VALUE ANALYSIS*

KENJI FUJIMOTO[†] AND JACQUELIEN M. A. SCHERPEN[‡]

Abstract. This paper discusses balanced realization and model order reduction for both continuous-time and discrete-time general nonlinear systems based on singular value analysis of the corresponding Hankel operators. Singular value analysis clarifies the gain structure of a given nonlinear operator. Here it is proved that singular value analysis of smooth Hankel operators defined on Hilbert spaces can be characterized by simple equations in terms of their states. A balanced realization and model order reduction procedure is derived based on it, and several important properties such as stability, balanced form, Hankel norm, controllability, and observability of the original system are preserved. The work improves the earlier results of [K. Fujimoto and J. M. A. Scherpen, *IEEE Trans. Automat. Control*, 50 (2005), pp. 2–18] and then continues with new balancing and model reduction results.

Key words. balanced realization, model reduction, singular value analysis, Hankel operators, nonlinear control systems

AMS subject classifications. 93B11, 93B20, 93C10

DOI. 10.1137/070695332

1. Introduction. In the theory of stable linear systems, the system Hankel operator plays an important role in a number of problems. Its relation to the state-space concept of balanced realizations, where the Hankel singular values are important, is well understood nowadays [32] and provides a powerful tool for model reduction of linear control systems. In this paper we propose a framework for a general class of stable nonlinear systems, where balanced realizations are directly related to the Hankel operator of the nonlinear system. This in turn provides a tool for model reduction of the nonlinear system, where properties such as stability, the Hankel norm, and the balanced form of the system are preserved. Our approach builds further upon the earlier developments in [6].

A first nonlinear extension of the linear state-space concept of balanced realizations has been introduced in [24], mainly based on studying the past input energy and the future output energy. Since then, many results on nonlinear state-space balancing, related minimality considerations, balancing near invariant manifolds, computational issues for model reduction, flow balancing, trajectory piecewise linear balancing, and empirical balancing for nonlinear systems have appeared in the literature; see, e.g., [7, 8, 10, 11, 12, 14, 16, 21, 23, 25, 27, 28, 30, 31].

In our earlier work, the relation of the state-space notion of balancing for finite dimensional, continuous-time, input affine nonlinear systems with the nonlinear Hankel operator has been considered; see, e.g., [11, 25, 26]. In particular, the *singular value functions* of [24], which can be viewed as a nonlinear state-space extension

*Received by the editors June 25, 2007; accepted for publication (in revised form) June 2, 2010; published electronically August 12, 2010. This work is partly supported by a visitor grant from NWO (Dutch Science Foundation).

<http://www.siam.org/journals/sicon/48-7/69533.html>

[†]Department of Mechanical Science and Engineering, Graduate School of Engineering, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, 464-8603, Japan (fujimoto@nagoya-u.jp).

[‡]Faculty of Mathematics and Natural Sciences, University of Groningen, ITM, Nijenborgh 4, 9747 AG Groningen, The Netherlands (J.M.A.Scherpen@rug.nl).

of the Hankel singular values in the linear case, can be related to nonlinear Hankel theory, [4, 6]. However, for obtaining the latter relation, a new characterization of (Hankel) singular value functions for nonlinear systems was proposed in [6], resulting in the definition of the so-called *axis singular value functions*. These functions have a close relationship to the gain structure of the Hankel operator and are characterized by *singular value analysis* [3] of the Hankel operator. Although the axis singular value functions are defined without using the state-space notion of singular value functions of [24], it was shown that they coincide at the coordinate axes when the system has a special state-space realization; hence the name *axis singular value functions*. In [5, 6], this special state-space realization was adopted and characterizes a nonlinear input-normal/output-diagonal realization. However, the latter realization only has a balance between the coordinate axes of the state space, whereas the balanced realization of a linear system also balances the relationship between the input-to-state behavior and the state-to-output behavior. From a realization and numerical point of view related to singular value decomposition, the latter property for linear systems is quite important [1].

A first objective of this paper is to generalize the main results of [6] to a larger class of nonlinear systems, as well as to provide a shorter and more elegant proof. With this, we establish a nonlinear singular value analysis directly related to the Hankel operator of finite and infinite dimensional, continuous- and discrete-time nonlinear systems. The proof is based upon nonlinear operators for this large class of nonlinear systems. The corresponding result of [6] is valid only for finite dimensional, continuous-time systems, and the proof, though constructive, is very long. The use of general nonlinear operators offers the possibility of shortening the proof and making it more insightful.

A second objective of this paper is to provide a truly balanced realization for nonlinear finite dimensional continuous- and discrete-time systems that offers a tool for model order reduction along the lines of the methods for linear systems. The starting point is the input-normal/output-diagonal realization that can be obtained almost immediately from the extended result mentioned above, but now restricted to finite dimensional nonlinear systems. From there, nonlinear balanced realizations are proposed. They provide a balance of the complete part of the state space that we consider, as opposed to providing a balance only among the coordinate axes as in [6]. The balancing method is applicable to discrete-time systems as well. Furthermore, the method offers a tool for the model reduction proposed in this paper, following along the lines of linear balanced model order reduction methods. It is shown that properties such as balanced form, axis singular value functions, stability, and the Hankel norm are preserved for the reduced order model obtained via the proposed model reduction procedures.

As mentioned above, we also propose balancing and order reduction methods for *discrete-time* nonlinear systems in this paper. So far, within our nonlinear balancing framework started in [24] for continuous-time input-affine systems, only characterizations and computations of the controllability and observability functions of discrete-time nonlinear systems have been reported in [17]. Typical nonlinear discrete-time systems are not input affine, and the earlier results of [6] for continuous-time input-affine systems are not directly applicable to discrete-time systems. Since the proposed approach of our current paper builds on nonlinear operators instead of state-space realizations, our new approach is also valid for discrete-time systems.

Our results basically build further on our earlier work in [24] and [6]. This means that our approach is valid in a neighborhood of an equilibrium point, and, depending

on the system, the neighborhood can be large or small. Other methods, such as originally presented in [28], and further developed in [27, 30, 31] for continuous-time systems and in [29] for discrete-time systems, are based on the flows of a system. These methods consider linearization around trajectories and use sliding time windows for the calculation of the reachability and observability Gramians. Then a return to the original nonlinear system is possible only for a limited class of systems. However, more generally, reachability and controllability Gramians in the sliding time window setting can be calculated approximately for the whole state space, thus yielding the basis for a balancing procedure of a large part of the state space. Nevertheless, relations with minimality and the Hankel operator are less clear than in our approach, among others, due to the approximation step in the flow balancing procedures. See [27, 28, 30, 31] for details. Furthermore, in [14] an approach based on the balancing method of [24] with polynomial approximations is treated by applying a balancing procedure to the different degrees of the polynomials separately. Also in [14] relations with minimality and the Hankel operator are less clear, with one of the reasons being the approximation step in the procedure.

Model order reduction based on balancing is a method based on singular value decompositions. However, there is also quite a bit of research effort being made in model order reduction methods based on Krylov methods and moment matching because of their computational advantages. See [1] for an overview for linear systems. Recently, a first extension of moment matching to the nonlinear case was obtained in [2]. Combinations with and relations to balancing are not yet developed.

The outline of this paper is as follows. Section 2 treats preliminaries and the problem setting. The linear systems case and the results of [6] are reviewed, and a nonlinear operator setting for Hankel analysis is introduced. Section 3 provides a singular value analysis of the Hankel operator, including an extension of a main result from [6] to general, finite/infinite dimensional, continuous-/discrete-time systems. It also provides a new elegant proof for the finite dimensional continuous-time result of [6]. This result is then used to determine balanced realizations for finite dimensional continuous- and discrete-time nonlinear systems in section 4. The developments of section 4 are then used in section 5 as a tool for model reduction based on the balanced realizations for both continuous- and discrete-time systems. It is shown that for continuous-time systems, balanced truncation is a suitable method for preserving certain balanced realization properties. However, for discrete-time systems, balanced truncation does not preserve these balancing properties; hence other order reduction strategies are considered. It is shown that order reduction based on singular perturbations analysis of the balanced realization does preserve the desired balanced realization properties. Finally, in section 6 we end with some conclusions.

Notation. The mathematical notation used throughout is fairly standard. If $x \in \mathbb{R}^n$, the norm is given as $\|x\| = (x^T x)^{1/2}$. If $x \in L_2[a, b]$, the norm is given as $\|x\| = (\int_a^b \|x(t)\|^2 dt)^{1/2}$. Similarly, for discrete-time signals if $x \in \ell_2[a, b]$, then the norm is given as $\|x\| = (\sum_{k=a}^b \|x(k)\|^2)^{1/2}$. Note that the type of norm is given by the space of the signal. The symbols \mathbb{R} and \mathbb{Z} denote the set of real numbers and the set of integers, respectively. Further, the half subsets of them are defined by $\mathbb{R}_+ := [0, \infty) \subset \mathbb{R}$, $\mathbb{R}_- := (-\infty, 0] \subset \mathbb{R}$, $\mathbb{Z}_+ := \{0, 1, 2, \dots\} \subset \mathbb{Z}$ and $\mathbb{Z}_- := \{0, -1, -2, \dots\} \subset \mathbb{Z}$, respectively. A condition about 0 means that this condition holds for a neighborhood of 0. Finally, $x(\pm\infty)$ is an abbreviation for $\lim_{t \rightarrow \pm\infty} x(t)$. Throughout this paper, by smooth we generally mean C^∞ , unless stated otherwise.

2. Preliminaries and problem setting. This section refers to the preliminary results on balanced realization for both linear and nonlinear systems and explains the problem setting for singular value analysis of nonlinear Hankel operators, which is the basic framework for balancing and model reduction for nonlinear control systems.

2.1. Linear systems as a paradigm. Here, we briefly review linear balancing theory; see, e.g., [32]. The presentation is such that the line of thinking in the nonlinear case is clarified. Consider a causal linear input-output system $\Sigma : L_2^m(\mathbb{R}_+) \rightarrow L_2^r(\mathbb{R}_+)$ with a state-space realization

$$(1) \quad u \mapsto y = \Sigma(u) : \begin{cases} \dot{x} &= Ax + Bu, \\ y &= Cx, \end{cases}$$

where $x(0) = 0$. The corresponding Hankel operator is given by the composition of the observability and controllability operators $\mathcal{H} = \mathcal{O} \circ \mathcal{C}$, where the observability and controllability operators, $\mathcal{O} : \mathbb{R}^n \rightarrow L_2^r(\mathbb{R}_+)$ and $\mathcal{C} : L_2^m(\mathbb{R}_+) \rightarrow \mathbb{R}^n$, respectively, are given by

$$\begin{aligned} x^0 \mapsto y &= \mathcal{O}(x^0) := Ce^{At}x^0, \\ u \mapsto x^0 &= \mathcal{C}(u) := \int_0^\infty e^{A\tau}Bu(\tau) \, d\tau. \end{aligned}$$

The Hankel, controllability, and observability operators are closely related to the observability and controllability Gramians, i.e., $Q = \mathcal{O}^* \circ \mathcal{O}$ and $P = \mathcal{C} \circ \mathcal{C}^*$. Furthermore, from, e.g., Theorem 8.1 in [32], we have the following relation.

THEOREM 1 (see [32]). *The operator $\mathcal{H}^* \circ \mathcal{H}$ and the matrix QP have the same nonzero eigenvalues.*

The square roots of the eigenvalues of QP are called the Hankel singular values of the system (1) and are denoted by σ_i 's where $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n$. They are the singular values of the Hankel operator, and the largest singular value equals the Hankel norm $\|\Sigma\|_H$ of the system Σ ,

$$(2) \quad \|\Sigma\|_H := \sup_{\substack{u \in L_2(\mathbb{R}_+) \\ u \neq 0}} \frac{\|\mathcal{H}(u)\|_{L_2}}{\|u\|_{L_2}} = \sigma_1.$$

Further, using a similarity transformation (linear coordinate transformation), we can diagonalize both P and Q so that

$$(3) \quad P = Q = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_n).$$

This state-space realization is called a *balanced realization*. The system is balanced in the following two senses:

- (i) P and Q are in a diagonal form; and
- (ii) $P = Q$, which means that the relationship between the input-to-state behavior and the state-to-output behavior is balanced.

Property (i) plays a central role in model reduction, and (ii) is important because it corresponds to a singular value decomposition that has certain numerical properties, and because it is important for realization algorithms; see, e.g., Chapter 4, and Chapter 7, Sections 7.3 and 7.4, of [1].

2.2. Hankel operators for nonlinear systems. In this paper, we consider a *Hankel operator* $\mathcal{H} : U \rightarrow Y$ for a nonlinear system defined on Hilbert spaces U and Y . Here, as in the linear case, we suppose that \mathcal{H} can be decomposed as

$$(4) \quad \mathcal{H} = \mathcal{O} \circ \mathcal{C}$$

with the *controllability operator* $\mathcal{C} : U \rightarrow X$ and the *observability operator* $\mathcal{O} : X \rightarrow Y$, where \mathcal{C} is surjective and X is also a Hilbert space. In the next examples, we study \mathcal{H} for particular dynamical systems. See [6] for the details.

Example 1. Consider an L_2 -stable finite dimensional continuous-time nonlinear system

$$(5) \quad \begin{cases} \dot{x} &= f(x, u, t), \\ y &= h(x, u, t). \end{cases}$$

The corresponding controllability operator $\mathcal{C} : L_2^m(\mathbb{R}_+) \rightarrow \mathbb{R}^n$ and observability operator $\mathcal{O} : \mathbb{R}^n \rightarrow L_2^p(\mathbb{R}_+)$ are defined by

$$(6) \quad x^0 = \mathcal{C}(u) : \begin{cases} \dot{x} &= -f(x, u, t), & x(\infty) = 0, \\ x^0 &= x(0), \end{cases}$$

$$(7) \quad y = \mathcal{O}(x^0) : \begin{cases} \dot{x}(t) &= f(x, 0, t), & x(0) = x^0, \\ y &= h(x, 0, t). \end{cases}$$

The Hankel operator is given by the composition (4) with $U = L_2^m(\mathbb{R}_+)$, $X = \mathbb{R}^n$ and $Y = L_2^p(\mathbb{R}_+)$.

Example 2. Consider an ℓ_2 -stable finite dimensional discrete-time nonlinear system

$$\begin{cases} x(t+1) &= f(x(t), u(t), t), \\ y(t) &= h(x(t), u(t), t). \end{cases}$$

Here we suppose that the $x(t+1) = f(x(t), u(t))$ is invertible with respect to $x(t)$. The corresponding controllability operator $\mathcal{C} : \ell_2^m(\mathbb{Z}_+) \rightarrow \mathbb{R}^n$ and observability operator $\mathcal{O} : \mathbb{R}^n \rightarrow \ell_2^p(\mathbb{Z}_+)$ are defined by

$$\begin{aligned} x^0 = \mathcal{C}(u) : & \begin{cases} x(t-1) &= f(x(t), u(t), t), & x(\infty) = 0, \\ x^0 &= x(0), \end{cases} \\ y = \mathcal{O}(x^0) : & \begin{cases} x(t+1) &= f(x(t), 0, t), & x(0) = x^0, \\ y(t) &= h(x(t), 0, t). \end{cases} \end{aligned}$$

The Hankel operator is given by the composition (4) with $U = \ell_2^m(\mathbb{Z}_+)$, $X = \mathbb{R}^n$ and $Y = \ell_2^p(\mathbb{Z}_+)$.

At this moment, we do not restrict ourselves to one of the above classes of systems, even though the results in [4, 6] are limited to finite dimensional continuous-time, time-invariant systems. Here, we study a much wider class of nonlinear systems including time-varying systems, input-nonaffine systems, and discrete-time systems.

The controllability and observability functions $L_c : X \rightarrow \mathbb{R}_+$ and $L_o : X \rightarrow \mathbb{R}_+$ with respect to the Hankel operator \mathcal{H} given in (4) are defined by

$$(8) \quad L_c(x^0) := \inf_{\substack{C(u)=x^0 \\ u \in U}} \frac{1}{2} \|u\|^2,$$

$$(9) \quad L_o(x^0) := \frac{1}{2} \|\mathcal{O}(x^0)\|^2.$$

If the pseudo-inverse $\mathcal{C}^\dagger : X \rightarrow U$ of $\mathcal{C} : U \rightarrow X$ defined by

$$(10) \quad \mathcal{C}^\dagger(x^0) := \arg \inf_{\substack{\mathcal{C}(u)=x^0 \\ u \in U}} \|u\|$$

exists, then L_c can be written as

$$(11) \quad L_c(x^0) = \frac{1}{2} \|\mathcal{C}^\dagger(x^0)\|^2.$$

In the linear case, we have the following relationship with the controllability and observability Gramians P and Q :

$$(12) \quad L_c(x^0) = \frac{1}{2} x^{0T} P^{-1} x^0,$$

$$(13) \quad L_o(x^0) = \frac{1}{2} x^{0T} Q x^0.$$

For a continuous-time input-affine nonlinear system of the form

$$\Sigma : \begin{cases} \dot{x} &= f(x) + g(x)u, \\ y &= h(x), \end{cases}$$

it is proved that the controllability and observability functions $L_c(x)$ and $L_o(x)$ are characterized by the solutions of a Hamilton–Jacobi equation (14) and a Lyapunov equation (15), [24]:

$$(14) \quad \frac{\partial L_c}{\partial x} f(x) + \frac{1}{2} \frac{\partial L_c}{\partial x} g(x) g(x)^T \frac{\partial L_c}{\partial x}^T = 0,$$

$$(15) \quad \frac{\partial L_o}{\partial x} f(x) + \frac{1}{2} h(x)^T h(x) = 0,$$

where 0 is an asymptotically stable equilibrium of $\dot{x} = -f - gg^T(\partial L_c(x)/\partial x)^T$ in a neighborhood of the origin. Furthermore, if the system is linear as in (1), and strict positivity of the solutions is assumed, then these partial differential equations reduce to the Lyapunov equations

$$\begin{aligned} AP + PA^T + B B^T &= 0, \\ QA + A^T Q + C^T C &= 0 \end{aligned}$$

for the Gramians P and Q as given in (12) and (13).

2.3. Singular value analysis of Hankel operators. For deriving a balanced realization of a given nonlinear system, we study the gain structure of the related Hankel operator \mathcal{H} , i.e., we examine

$$(16) \quad \sigma_{\max}(c) := \sup_{\|u\|=c} \frac{\|\mathcal{H}(u)\|}{\|u\|},$$

$$(17) \quad v_{\max}(c) := \arg \sup_{\|u\|=c} \frac{\|\mathcal{H}(u)\|}{\|u\|}.$$

Here we assume the existence of $v_{\max} \in U$. We add the constraint $\|u\| = c > 0$ because we are interested in the maximizing input for each input magnitude c .

Here we suppose that the Hankel operator \mathcal{H} is (Fréchet) differentiable.¹ Since v_{\max} defined in (17) is a critical point of $(\|\mathcal{H}(u)\|/\|u\|)$, $u = v_{\max}$ needs to satisfy

$$(18) \quad d\left(\frac{\|\mathcal{H}(u)\|}{\|u\|}\right)(du) = 0 \quad \text{s.t.} \quad \|u\| = c.$$

Doing the derivation of the first equation in (18), we obtain

$$\begin{aligned} 0 &= d\left(\frac{\|\mathcal{H}(u)\|}{\|u\|}\right)(du) \\ &= \frac{\|u\| \cdot d(\|\mathcal{H}(u)\|)(du) - \|\mathcal{H}(u)\| \cdot d(\|u\|)(du)}{\|u\|^2} \\ &= \frac{(\|u\|/\|\mathcal{H}(u)\|)\langle \mathcal{H}(u), d\mathcal{H}(u)(du) \rangle - (\|\mathcal{H}(u)\|/\|u\|)\langle u, du \rangle}{\|u\|^2} \\ &= \frac{(\|u\|/\|\mathcal{H}(u)\|)\langle (d\mathcal{H}(u))^* \circ \mathcal{H}(u), du \rangle - (\|\mathcal{H}(u)\|/\|u\|)\langle u, du \rangle}{\|u\|^2} \\ (19) \quad &= \frac{\langle (d\mathcal{H}(u))^* \circ \mathcal{H}(u) - (\|\mathcal{H}(u)\|/\|u\|)^2 u, du \rangle}{\|u\| \cdot \|\mathcal{H}(u)\|}. \end{aligned}$$

On the other hand, differentiating the constraint in (18) reduces to

$$\langle u, du \rangle = 0.$$

Hence we can rewrite the problem (18) as

$$\langle (d\mathcal{H}(u))^* \circ \mathcal{H}(u) - (\|\mathcal{H}(u)\|/\|u\|)^2 u, du \rangle = 0 \quad \forall du \quad \text{s.t.} \quad \langle u, du \rangle = 0.$$

Finally, we obtain an alternative formulation of (18) as follows:

$$(20) \quad (d\mathcal{H}(u))^* \circ \mathcal{H}(u) = \lambda u, \quad \lambda \in \mathbb{R}.$$

Equation (20) characterizes all critical inputs u as well as the maximizing input v_{\max} . Note that this equation no longer contains the parameter c . This means that the solutions to this equation will be implicitly parameterized by the parameter c . Essentially, this fact implies that the solution set is made up of curves in the input signal space which characterize the coordinate axes of the balanced coordinates. Then, consequently, we can obtain the nonlinear balanced realization. In order to characterize the balanced realization, we are interested in characterizing the states (at $t = 0$) achieving the critical points of $\|\mathcal{H}(u)\|/\|u\|$. Hence it is natural to restrict our problem (20) to a subset $\text{Im}\mathcal{C}^\dagger$ of the input signal space U since its elements have one-to-one correspondence to those of the state space X . Therefore we will solve (20) with

$$(21) \quad u \in \text{Im}\mathcal{C}^\dagger$$

in what follows. Let us define the solutions for u in the above equations (20) and (21) by v . We call investigation of the solutions v and λ for the above equations *singular*

¹Here the operator $d(\cdot)$ denotes the Fréchet derivative. The Fréchet derivative df of a given function $f: X \rightarrow Y$ with Banach spaces X and Y satisfies

$$f(x + \xi) - f(x) = df(x)(\xi) + o(\|\xi\|_X),$$

and $df(x)(\xi)$ is linear in ξ .

value analysis of \mathcal{H} . Singular value analysis proposed here was called “differential eigenstructure of Hankel operators” in the authors’ former paper [6]. It should be noted that the singular vector v is an eigenvector of the operator $(d\mathcal{H}(u))^* \circ \mathcal{H}(u)$. The vector v is a *singular vector*, and the corresponding scalar σ defined by

$$(22) \quad \sigma(v) := \frac{\|\mathcal{H}(v)\|}{\|v\|}$$

is called a *singular value* of \mathcal{H} . See [3] for the details of singular value analysis of nonlinear operators.

Furthermore, it also follows from (19) that the critical points of $\|\mathcal{H}(u)\|/\|u\|$ without the constraint $\|u\| = c$,

$$(23) \quad d\left(\frac{\|\mathcal{H}(u)\|}{\|u\|}\right)(du) = 0,$$

are characterized by

$$(24) \quad (d\mathcal{H}(u))^* \circ \mathcal{H}(u) = \sigma^2 u$$

with the singular value σ as defined in (22). Namely, λ in (20) coincides with the square of the singular value σ^2 at the critical points of $\|\mathcal{H}(u)\|/\|u\|$ without constraint.

2.4. Nonlinear input-normal realization. This section briefly reviews the authors’ preliminary results on input-normal/output-diagonal realizations for time-invariant input-affine nonlinear systems reported in [6].

Consider a smooth input-affine nonlinear system Σ ,

$$(25) \quad u \mapsto y = \Sigma(u) : \begin{cases} \dot{x} &= f(x) + g(x)u, \\ y &= h(x), \end{cases}$$

with $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$, and $y(t) \in \mathbb{R}^r$, with $x = 0$ an equilibrium point for $u = 0$. We introduce the following technical assumptions in order to state the main results of [6].

Assumption B1. Suppose that the system Σ in (25) is asymptotically stable about the origin and that there exists a neighborhood of the origin where the operators \mathcal{O} , \mathcal{C} , and \mathcal{C}^\dagger exist and are smooth.

Assumption B2. Suppose that the Hankel singular values of the Jacobian linearization of the system Σ around $x = 0$ are nonzero and distinct.

The nonlinear state-space developments of [24] give an input-normal/output-diagonal form of system (25) as follows.

THEOREM 2 (see [24]). *Consider the operator Σ with the asymptotically stable state-space realization (25). Suppose that Assumptions B1 and B2 hold. Then there exist a neighborhood W of the origin and a smooth coordinate transformation $x = \Phi(z)$ on W converting Σ into an input-normal/output-diagonal form, where*

$$(26) \quad L_c(\Phi(z)) = \frac{1}{2} z^T z,$$

$$(27) \quad L_o(\Phi(z)) = \frac{1}{2} \sum_{i=1}^n z_i^2 \tau_i(z),$$

with $\tau_1(z) \geq \dots \geq \tau_n(z)$ being the so-called smooth singular value functions on W .

The above *input-normal/output-diagonal realization* is important for what follows. However, this result is incomplete in the sense that properties (i) and (ii) explained below equation (3) are not exactly fulfilled, as explained in section 4. Indeed, this realization and the τ_i functions are not unique [11], and, consequently, the corresponding model reduction procedure gives different reduced models according to the choices of different sets of singular value functions. In [6] this issue is tackled by considering the relation with the Hankel operator, and a more precise input-output characterization for the input-normal/output-diagonal realization is given.

The Hankel operator for the system (5) is given as in Example 1, where we restrict the plant system to be input affine. Instead of considering the eigenstructure of $\mathcal{H}^* \circ \mathcal{H}$ as in the linear case given in Theorem 1, we consider the solution pair $\lambda \in \mathbb{R}$ and $v \in L_2(\mathbb{R}_+)$ of the *singular value analysis* of \mathcal{H} characterized by (20) and (21). Using the latter singular value analysis, we have the following result.

THEOREM 3 (see [6]). *Consider the Hankel operator \mathcal{H} in (4). Suppose that Assumptions B1 and B2 hold. Then there exists a neighborhood $S_0 \subset \mathbb{R}$ of 0, n smooth functions $\sigma_i : S_0 \rightarrow \mathbb{R}_+$, $i \in \{1, 2, \dots, n\}$, such that*

$$\min\{\sigma_i(s), \sigma_i(-s)\} \geq \max\{\sigma_{i+1}(s), \sigma_{i+1}(-s)\}$$

holds for all $s \in S_0$ and all $i \in \{1, 2, \dots, n-1\}$ and such that there exist n distinct smooth curves $\xi_i : S_0 \rightarrow X$ satisfying $\xi_i(0) = 0$ and

$$\begin{aligned} L_c(\xi_i(s)) &= \frac{s^2}{2}, & L_o(\xi_i(s)) &= \frac{\sigma_i^2(s)}{2} s^2, \\ \frac{\partial L_o}{\partial x}(\xi_i(s)) &= \lambda_i(s) \frac{\partial L_c}{\partial x}(\xi_i(s)) \end{aligned}$$

with

$$\lambda_i(s) := \sigma_i^2(s) + \frac{s}{2} \frac{d\sigma_i^2(s)}{ds}.$$

In particular, if $S_0 = \mathbb{R}$, then

$$\sup_{\substack{u \in U \\ u \neq 0}} \frac{\|\mathcal{H}(u)\|}{\|u\|} = \sup_{s \in S_0} \sigma_1(s).$$

Here the parameter c in (18) is given by $c = |s|$ with scalar parameter s parameterizing the solutions in the theorem. Furthermore, based on the above theorem, a more precise version of the input-normal/output-diagonal realization was derived.

THEOREM 4 (see [6, Theorem 8]). *Consider the operator Σ with the state-space realization (25). Suppose that Assumptions B1 and B2 hold. Then there exist a neighborhood W of 0 and a coordinate transformation $x = \Phi(z)$ on W converting the system into an input-normal form (26), (27) satisfying the properties*

$$(28) \quad z_i = 0 \Leftrightarrow \frac{\partial L_c(\Phi(z))}{\partial z_i} = 0 \Leftrightarrow \frac{\partial L_o(\Phi(z))}{\partial z_i} = 0$$

for all $i \in \{1, 2, \dots, n\}$ on W . Furthermore

$$\begin{aligned} \tau_i(0, \dots, 0, \underbrace{z_i}_{i\text{th}}, 0, \dots, 0) &= \sigma_i(z_i)^2, \\ \frac{\partial \tau_i}{\partial z}(0, \dots, 0, \underbrace{z_i}_{i\text{th}}, 0, \dots, 0) &= (0, \dots, 0, \underbrace{\frac{d\sigma_i(z_i)^2}{dz_i}}_{i\text{th}}, 0, \dots, 0) \end{aligned}$$

holds for all $i \in \{1, 2, \dots, n\}$. In particular, if $W = \mathbb{R}^n$, then

$$\|\Sigma\|_H^2 = \sup_{z_1 \in \mathbb{R}} \tau_1(z_1, 0, \dots, 0).$$

By this theorem, we can obtain an input-normal/output-diagonal realization which has a close relationship to the Hankel operator. In fact, this theorem gives the nonlinear version of property (i) explained below equation (3). However, the nonlinear version of characterization (ii) has not been obtained so far. This is one of the problems considered in the remainder of the present paper. It is also noted that Theorems 3 and 4 are only for continuous-time input-affine nonlinear systems, and their proofs given in [6] are quite long. In the following section, we derive the result for a wider class of (finite dimensional) nonlinear systems, with a much simpler proof based on the analysis of Hankel operator \mathcal{H} .

3. Singular value analysis and observability and controllability functions. The objective of this section is to relate the singular value analysis of the Hankel operator \mathcal{H} of (20) and (21) for a general class of nonlinear systems to the controllability and observability operators independent of the state-space representation. In this section we extend the result of Lemma 4 of [6] to the general class of systems including both finite and infinite dimensional, and continuous-time and discrete-time nonlinear systems. At the same time, the extension provides a new, briefer, and more elegant proof than in [6] for the existing result for finite dimensional, continuous-time, input-affine systems given by Theorem 4 in this paper.

Let us consider a system with an input signal space U , an output signal space Y , and a state space X with the controllability and observability operators $\mathcal{C} : U \rightarrow X$ and $\mathcal{O} : X \rightarrow Y$. The Hankel operator \mathcal{H} is defined by (4).

Assumption A1. The operators $\mathcal{C} : U \rightarrow X$, $\mathcal{O} : X \rightarrow Y$ and $\mathcal{C}^\dagger : X \rightarrow U$ exist and are differentiable.

Under this assumption, we can obtain an alternative characterization of singular value analysis of the Hankel operator on the signal space X , i.e., different from (20).

THEOREM 5. *Suppose that Assumption A1 holds. Assume moreover that there exist $\lambda \in \mathbb{R}$ and $\xi \in X$ satisfying*

$$(29) \quad dL_o(\xi) = \lambda dL_c(\xi).$$

Then $u = v \in U$ defined by

$$(30) \quad v := \mathcal{C}^\dagger(\xi)$$

satisfies (20) and (21), i.e., the equations for singular value analysis of \mathcal{H} .

Proof. As preparation for the proof of the theorem, we need to clarify some properties of the signal space $\text{Im}\mathcal{C}^\dagger$ given in (21). By Assumption A1, both \mathcal{C} and \mathcal{C}^\dagger exist and are differentiable. Hence the constraint (21) can be characterized by singular value analysis of $\mathcal{C}^\dagger \circ \mathcal{C}$ since $\mathcal{C}^\dagger \circ \mathcal{C}$ is a projector onto $\text{Im}\mathcal{C}^\dagger$. That is, any element in $\text{Im}\mathcal{C}^\dagger$ is characterized by

$$(31) \quad \arg \sup_{u \in U} \frac{\|\mathcal{C}^\dagger \circ \mathcal{C}(u)\|}{\|u\|}$$

with the maximum singular value 1, since

$$\frac{\|\mathcal{C}^\dagger \circ \mathcal{C}(u)\|}{\|u\|} = 1 \quad u \in \text{Im } \mathcal{C}^\dagger,$$

$$\frac{\|\mathcal{C}^\dagger \circ \mathcal{C}(u)\|}{\|u\|} < 1 \quad \text{otherwise}$$

hold for the definition of \mathcal{C}^\dagger in (10). Therefore, any element $v \in \text{Im } \mathcal{C}^\dagger$ satisfies the critical points condition without constraint as in (24) with the (maximum) singular value $\sigma = \|\mathcal{C}^\dagger \circ \mathcal{C}(v)\|/\|v\| = 1$, i.e.,

$$(\text{d}(\mathcal{C}^\dagger \circ \mathcal{C})(v))^* \circ (\mathcal{C}^\dagger \circ \mathcal{C}(v)) = 1^2 \cdot v,$$

which reduces to

$$(32) \quad (\text{d}\mathcal{C}(v))^* \circ (\text{d}\mathcal{C}^\dagger(\mathcal{C}(v)))^* \circ \mathcal{C}^\dagger \circ \mathcal{C}(v) = v,$$

where v is a singular vector.

Now we can prove the theorem using (32). Suppose that there exist $\lambda \in \mathbb{R}$ and $\xi \in X$ satisfying (29) and define the corresponding input $v \in U$ by (30). Then v is an element of $\text{Im } \mathcal{C}^\dagger$ by its definition, so (32) holds with the signal v thus defined. Substituting L_o and L_c in (9) and (11) for (29) yields

$$(33) \quad (\text{d}\mathcal{O}(\xi))^* \circ (\mathcal{O}(\xi)) = \lambda (\text{d}\mathcal{C}^\dagger(\xi))^* \circ (\mathcal{C}^\dagger(\xi)),$$

since

$$(34) \quad \text{d}L_c(x)(\text{d}x) = \langle \mathcal{C}^\dagger(x), \text{d}\mathcal{C}^\dagger(x)(\text{d}x) \rangle = \langle (\text{d}\mathcal{C}^\dagger(x))^* \circ (\mathcal{C}^\dagger(x)), \text{d}x \rangle,$$

$$(35) \quad \text{d}L_o(x)(\text{d}x) = \langle \mathcal{O}(x), \text{d}\mathcal{O}(x)(\text{d}x) \rangle = \langle (\text{d}\mathcal{O}(x))^* \circ (\mathcal{O}(x)), \text{d}x \rangle.$$

Due to the definition of v ,

$$(36) \quad \xi = \mathcal{C}(v)$$

holds. Substitute this for (33); then we obtain

$$(37) \quad (\text{d}\mathcal{C}^\dagger(\mathcal{C}(v)))^* \circ \mathcal{C}^\dagger \circ \mathcal{C}(v) = \frac{1}{\lambda} (\text{d}\mathcal{O}(\mathcal{C}(v)))^* \circ \mathcal{O} \circ \mathcal{C}(v).$$

Then, by further substituting this equation for (32), we have

$$(\text{d}\mathcal{C}(v))^* \circ \left(\frac{1}{\lambda} (\text{d}\mathcal{O}(\mathcal{C}(v)))^* \circ \mathcal{O} \circ \mathcal{C}(v) \right) = v.$$

Since $(\text{d}\mathcal{C}(v))^*$ is a linear operator, this reduces to

$$(38) \quad (\text{d}\mathcal{C}(v))^* \circ (\text{d}\mathcal{O}(\mathcal{C}(v)))^* \circ \mathcal{O} \circ \mathcal{C}(v) = \lambda v.$$

On the other hand, substituting \mathcal{H} in (4) for (20) yields

$$(39) \quad (\text{d}\mathcal{C}(v))^* \circ (\text{d}\mathcal{O}(\mathcal{C}(v)))^* \circ \mathcal{O} \circ \mathcal{C}(v) = \lambda v,$$

which coincides with (38). Hence the input v defined by (30) satisfies (20), which is the equation for singular value analysis. Also (21) trivially follows from the definition of v in (30). This proves the theorem. \square

This theorem gives us a sufficient condition for singular value analysis of \mathcal{H} characterized in (20) and (21) for a nonlinear system of which the state-space is not yet specified. Condition (29) is easier to check than (20) and (21) if the dimension of the intermediate state-space X is smaller than that of the input signal space U , e.g., $U = L_2$ and $X = \mathbb{R}^n$.

Note that the corresponding singular value σ defined in (22) is given by

$$\begin{aligned}\sigma(\mathcal{C}^\dagger(\xi)) &= \frac{\|\mathcal{H}(\mathcal{C}^\dagger(\xi))\|}{\|\mathcal{C}^\dagger(\xi)\|} = \frac{\|\mathcal{O} \circ \mathcal{C} \circ \mathcal{C}^\dagger(\xi)\|}{\|\mathcal{C}^\dagger(\xi)\|} = \frac{\|\mathcal{O}(\xi)\|}{\|\mathcal{C}^\dagger(\xi)\|} \\ &= \sqrt{\frac{(1/2)\|\mathcal{O}(\xi)\|^2}{(1/2)\|\mathcal{C}^\dagger(\xi)\|^2}} = \sqrt{\frac{L_o(\xi)}{L_c(\xi)}}.\end{aligned}$$

In particular, if we can characterize all ξ_i 's of (29) and let σ_i 's denote the corresponding singular values, then clearly we can obtain the Hankel norm, which is the gain of the Hankel operator, in the following way:

$$\sup_{\substack{u \in U \\ u \neq 0}} \frac{\|\mathcal{H}(u)\|}{\|u\|} = \max_i \sup_{\xi_i \neq 0} \sigma_i(\mathcal{C}^\dagger(\xi_i)).$$

Example 3. Suppose that our plant system is the linear dynamical system given in section 2.1. Then (29) yields

$$\xi^T Q = \lambda \xi^T P^{-1}$$

with the controllability and observability Gramians P and Q , which is equivalent to

$$PQ \xi = \lambda \xi.$$

That is, ξ is the eigenvector of PQ , and all eigenvectors of PQ form the basis for the balanced realization; i.e., after the balancing transformation the eigenvectors are transformed into the new coordinate axes of the system. Furthermore, $\lambda = \sigma^2$, where the σ 's are the Hankel singular values.

Example 4. Suppose that our plant system is the dynamical system given in Example 1 or 2. Then the solution of singular value analysis of the corresponding Hankel operator can be characterized by an algebraic equation

$$(40) \quad \frac{\partial L_o}{\partial x}(\xi) = \lambda \frac{\partial L_c}{\partial x}(\xi).$$

In comparison to the linear case mentioned above in Example 3, the set of ξ 's plays the role of the eigenvectors, and thus they can be viewed as the axes of the balanced coordinates.

Note that we do not require any state-space realization of the operators here. Hence, Theorem 5 is applicable to very general nonlinear systems, including both continuous- and discrete-time, finite and infinite dimensional, and input-affine and input-nonaffine dynamical systems.

4. Balanced realization. We now study balanced realizations based on the solutions of (29), that is, balanced realizations whose coordinate axes coincide with the ξ 's. The result on the singular value analysis of Hankel operators given in Theorem 5 holds with any nonlinear system such as continuous-time and discrete-time

systems. This allows one to obtain the balanced realization of both continuous-time and discrete-time input-nonaffine nonlinear systems. Here, we first extend the input-normal/output-diagonal balancing procedure given in section 2 to more general systems, and then we study balanced realizations based on the latter extended procedure. We now restrict ourselves to systems with a finite dimensional state space.

4.1. Input-normal/output-diagonal balancing. In order to generalize Theorems 3 and 4 to general systems with finite dimensional state space, we need to employ the following assumption.

Assumption A2. Suppose that $X \subset \mathbb{R}^n$, that $(\partial^2 L_c(x)/\partial x^2)(0)$ and $(\partial^2 L_o(x)/\partial x^2)(0)$ are positive definite, and that the eigenvalues of $((\partial^2 L_c/\partial x^2)(0))^{-1}((\partial^2 L_o/\partial x^2)(0))$ are distinct.

Under Assumption A2, we can prove the existence of n independent $\xi_i, i = 1, \dots, n$, solutions of (29) (or (40)).

THEOREM 6. *Consider a nonlinear system with a Hankel operator \mathcal{H} as in (4). Suppose that Assumptions A1 and A2 hold. Then the statements of Theorem 3 hold.*

Proof. The proof of Theorem 3 does not use a specific state-space realization of the plant and it depends only on Assumption B2, which can be trivially replaced by Assumption A2. This proves the theorem. \square

This result can be used for input-normal balanced realization for both continuous-time and discrete-time nonlinear systems as follows.

THEOREM 7. *Consider a nonlinear system with a Hankel operator \mathcal{H} as in (4). Suppose that Assumptions A1 and A2 hold. Then the statements of Theorem 4 hold.*

Proof. The proof follows along the same lines as the proof of Theorem 5. \square

Hence, if we assume that the system has a finite-dimensional state space, then the input-normal/output-diagonalization procedure given in Theorems 6 and 7 is applicable to the finite-dimensional continuous-time and discrete-time state-space systems as given in Examples 1 and 2. Note that, in contrast to the long, and less elegant, proofs of the continuous time result in [6], now we do not require explicit descriptions of the dynamics of the system.

4.2. Balanced realization. So far, we have focused on extending results from [6] to a more general class of systems. The obtained input-normal/output-diagonal representation fulfills item (i) given below equation (3). However, item (ii) given below equation (3) is not fulfilled. For a truly balanced realization, item (ii) should also be fulfilled. In this section we propose a *new* truly balanced realization in the sense that item (ii) is also fulfilled. The proposed realization clarifies the input-to-state and state-to-output behavior of nonlinear dynamical systems, i.e., the relationship given in item (ii) below equation (3). We first obtain a new input-normal/output-diagonal form for a 2-dimensional system. The latter form plays a key role in obtaining a balanced form for a finite dimensional systems of order n .

LEMMA 1. *Consider a nonlinear system with a Hankel operator \mathcal{H} as in (4), fulfilling Assumptions A1 and A2. Furthermore, assume that $X \subset \mathbb{R}^2$. Then there exist a neighborhood $W \subset X$ of the origin and a coordinate transformation $x = \Phi(z)$ on W converting the system into the form*

$$\begin{aligned} L_c(\Phi(z)) &= \frac{1}{2} (z_1^2 + z_2^2), \\ L_o(\Phi(z)) &= \frac{1}{2} ((z_1 \sigma_1(z_1))^2 + (z_2 \sigma_2(z_2))^2). \end{aligned}$$

Proof. Since the proof is long and technical, we defer it to the appendix. \square

Using this lemma recursively and repeatedly along similar lines as the proof of Theorem 8 in [6], a new input-normal/output-diagonal realization can be obtained. In the new realization, all the coordinate axes of the state-space appear decoupled in the observability and controllability functions.

THEOREM 8. *Consider a nonlinear operator with a Hankel operator \mathcal{H} as in (4). Suppose that Assumptions A1 and A2 hold. Then there exist a neighborhood $W \subset X$ of the origin and a coordinate transformation $x = \Phi(z)$ on W , converting the system into the form*

$$(41) \quad L_c(\Phi(z)) = \frac{1}{2} z^T z,$$

$$(42) \quad L_o(\Phi(z)) = \frac{1}{2} \sum_{i=1}^n (z_i \sigma_i(z_i))^2.$$

In particular, if $W = X$, then

$$\|\Sigma\|_H = \sup_{\substack{z_1 \\ \Phi(z_1, 0, \dots, 0) \in X}} \sigma_1(z_1).$$

Proof. See the appendix. \square

Example 5. Let us take an example from [6]. Consider a nonlinear system

$$(43) \quad \begin{cases} \dot{x} &= f(x) + g(x)u, \\ y &= h(x) \end{cases}$$

with

$$\begin{aligned} f(x) &= \begin{pmatrix} -9x_1 - x_1^5 - 2x_1^3x_2^2 - x_1x_2^4 \\ -9x_2 - x_1^4x_2 - 2x_1^2x_2^3 - x_2^5 \end{pmatrix}, \\ g(x) &= \begin{pmatrix} \sqrt{18 + 2x_1^4 + 4x_1^2x_2^2 + 2x_2^4} & 0 \\ 0 & \sqrt{18 + 2x_1^4 + 4x_1^2x_2^2 + 2x_2^4} \end{pmatrix}, \\ h(x) &= \begin{pmatrix} \frac{(6x_1 - 2x_1^5 - 4x_1^3x_2^2 - 2x_1x_2^4)\sqrt{18 + 2x_1^4 + 4x_1^2x_2^2 + 2x_2^4}}{1 + x_1^4 + 2x_1^2x_2^2 + x_2^4} \\ \frac{(3x_2 - x_1^4x_2 - 2x_1^2x_2^3 - x_2^5)\sqrt{18 + 2x_1^4 + 4x_1^2x_2^2 + 2x_2^4}}{1 + x_1^4 + 2x_1^2x_2^2 + x_2^4} \end{pmatrix}. \end{aligned}$$

It is shown in [6] that this system is balanced in the sense of Theorems 4 and 7; that is, its controllability and observability functions $L_c(x)$ and $L_o(x)$ are

$$(44) \quad L_c(x) = \frac{1}{2} x^T x,$$

$$(45) \quad L_o(x) = \frac{1}{2} (x_1^2 \tau_1(x) + x_2^2 \tau_2(x))$$

with

$$(46) \quad \tau_1(x) = \frac{4(9 + x_1^4 + 2x_1^2x_2^2 + x_2^4)}{1 + x_1^4 + 2x_1^2x_2^2 + x_2^4} = 4(9 - 8x_1^4 - 16x_1^2x_2^2 - 8x_2^4) + o(\|x\|^7),$$

$$(47) \quad \tau_2(x) = \frac{9 + x_1^4 + 2x_1^2x_2^2 + x_2^4}{1 + x_1^4 + 2x_1^2x_2^2 + x_2^4} = 9 - 8x_1^4 - 16x_1^2x_2^2 - 8x_2^4 + o(\|x\|^7).$$

They satisfy the relationship (28). Since the dimension of the system is 2, the coordinate transformation obtaining the input-normal form characterized in Theorem 8 is

given by $x = \Theta^{-1} \circ \Psi \circ \Theta(z)$ in the proof of Lemma 1. In order to compute it explicitly, we employ the Taylor series approximation up to the 5th order in a similar way to [8]. First of all, it follows from Theorems 4 and 7 that the singular value functions $\sigma_i(\cdot)$'s are given by

$$(48) \quad \sigma_1(x_1) = \sqrt{\tau_1(x_1, 0)} = 2\sqrt{\frac{9+x_1^4}{1+x_1^4}} = 6 - \frac{8}{3}x_1^4 + o(x_1^7),$$

$$(49) \quad \sigma_2(x_2) = \sqrt{\tau_2(0, x_2)} = \sqrt{\frac{9+x_2^4}{1+x_2^4}} = 3 - \frac{4}{3}x_2^4 + o(x_2^7).$$

Therefore, the balanced realization with the state z in Theorem 8 should have the controllability and observability functions

$$(50) \quad L_c(\Phi(z)) = \frac{1}{2}z^T z + o(\|z\|^6),$$

$$(51) \quad L_o(\Phi(z)) = \frac{1}{2}((z_1\sigma_1(z_1))^2 + (z_2\sigma_2(z_2))^2) + o(\|z\|^6).$$

Substituting (44)–(49) for (50) and (51), we obtain a pair of equations for the coordinate function $\Phi(z) = (\phi_1(z), \phi_2(z))$ as follows:

$$\begin{aligned} \phi_1(z)^2 + \phi_2(z)^2 &= z_1^2 + z_2^2 + o(\|z\|^6), \\ (9 - 8\phi_1(z)^4 - 16\phi_1(z)^2\phi_2(z)^2 - 8\phi_2(z)^4)(\phi_1(z)^2 + 4\phi_2(z)^2) \\ &= \left(6z_1 - \frac{8}{3}z_1^3\right)^2 + \left(3z_2 - \frac{4}{3}z_2^3\right)^2 + o(\|z\|^6). \end{aligned}$$

Solving the above equations for $\phi_1(z)$ and $\phi_2(z)$ with the Taylor series approximation up to the order 5, we obtain the solution

$$x = \Phi(z) = \begin{pmatrix} \phi_1(z) \\ \phi_2(z) \end{pmatrix} = \begin{pmatrix} z_1 + \frac{4}{3}z_1^3z_2^2 + \frac{8}{9}z_1z_2^4 \\ z_2 - \frac{4}{3}z_1^4z_2 - \frac{8}{9}z_1^2z_2^3 \end{pmatrix} + o(\|z\|^5).$$

The transformed system is described in the coordinate z by

$$(52) \quad \begin{cases} \dot{z} &= \hat{f}(z) + \hat{g}(z)u, \\ y &= \hat{h}(z) \end{cases}$$

with

$$\begin{aligned} \hat{f}(z) &= \begin{pmatrix} -9z_1 - z_1^5 + 46z_1^3z_2^2 + 31z_1z_2^4 \\ -9z_2 - 49z_1^4z_2 - 34z_1^2z_2^3 - z_2^5 \end{pmatrix} + o(\|z\|^5), \\ \hat{g}(z) &= \begin{pmatrix} 3\sqrt{2} + \frac{\sqrt{2}}{6}z_1^4 - \frac{35\sqrt{2}}{3}z_1^2z_2^2 - \frac{5\sqrt{2}}{2}z_2^4 & -8\sqrt{2}z_1^3z_2 - \frac{32\sqrt{2}}{3}z_1z_2^3 \\ -8\sqrt{2}z_1^3z_2 - \frac{32\sqrt{2}}{3}z_1z_2^3 & 3\sqrt{2} + \frac{25\sqrt{2}}{6}z_1^4 + \frac{25\sqrt{2}}{3}z_1^2z_2^2 + \frac{\sqrt{2}}{6}z_2^4 \end{pmatrix} \\ &\quad + o(\|z\|^4), \end{aligned}$$

$$\hat{h}(z) = \begin{pmatrix} 18\sqrt{2}z_1 - 23\sqrt{2}z_1^5 - 22\sqrt{2}z_1^3z_2^2 - 7\sqrt{2}z_1z_2^4 \\ 9\sqrt{2}z_2 - \frac{47\sqrt{2}}{2}z_1^4z_2 - 31\sqrt{2}z_1^2z_2^3 - \frac{23\sqrt{2}}{2}z_2^5 \end{pmatrix} + o(\|z\|^5).$$

It is easy to verify that this system has the controllability and observability functions (50) and (51); that is, it is balanced in the sense of Theorem 8.

Once we obtain the observability and controllability functions which are decoupled on the coordinate axes, it is easy to obtain the balanced realization, i.e., a realization with a balance between the input-to-state behavior and the state-to-output behavior.

THEOREM 9. *Consider a nonlinear system with a Hankel operator \mathcal{H} as in (4). Suppose that Assumptions A1 and A2 hold. Then there exist a neighborhood W of the origin and a coordinate transformation $x = \bar{\Phi}(\bar{z})$ on W converting the system into the form*

$$(53) \quad L_c(\bar{\Phi}(\bar{z})) = \frac{1}{2} \sum_{i=1}^n \frac{\bar{z}_i^2}{\bar{\sigma}_i(\bar{z}_i)},$$

$$(54) \quad L_o(\bar{\Phi}(\bar{z})) = \frac{1}{2} \sum_{i=1}^n \bar{z}_i^2 \bar{\sigma}_i(\bar{z}_i).$$

In particular, if $W = X$, then

$$\|\Sigma\|_H = \sup_{\substack{\bar{z}_1 \\ \bar{\Phi}(\bar{z}_1, 0, \dots, 0) \in X}} \bar{\sigma}_1(\bar{z}_1).$$

Proof. First of all, let us apply the coordinate transformation of Theorem 8 to obtain the z coordinate. Next we apply another coordinate transformation $z = \Phi^{-1} \circ \bar{\Phi}(\bar{z}) = (\bar{\phi}_1(\bar{z}_1), \bar{\phi}_2(\bar{z}_2), \dots, \bar{\phi}_n(\bar{z}_n))$ with $\bar{z}_i = \bar{\phi}_i^{-1}(z_i) := z_i \sqrt{\sigma_i(z_i)}$ to this system. Then we obtain a state-space realization with the controllability and observability functions as in (53) and (54) with

$$\bar{\sigma}_i(\bar{z}_i) := \sigma_i(\bar{\phi}_i(\bar{z}_i)),$$

which proves the theorem. \square

We call the state-space realization described in the new coordinates z in Theorem 9 a *balanced realization* of the given nonlinear system. In fact, the controllability and observability functions can be rewritten as

$$\begin{aligned} L_c(\bar{\Phi}(\bar{z})) &= \frac{1}{2} \bar{z}^T P(\bar{z})^{-1} \bar{z}, \quad L_o(\bar{\Phi}(\bar{z})) = \frac{1}{2} \bar{z}^T Q(\bar{z}) \bar{z}, \\ P(\bar{z}) &= Q(\bar{z}) = \text{diag}(\bar{\sigma}_1(\bar{z}_1), \bar{\sigma}_2(\bar{z}_2), \dots, \bar{\sigma}_n(\bar{z}_n)), \end{aligned}$$

which are very similar to those of the linear balanced realization (3). The coordinates could now be called “uncorrelated,” a terminology that was previously used in, e.g., [4, 28]. The $\bar{\sigma}_i$ and σ_i functions have the same value and are the singular values of the Hankel operator \mathcal{H} . Furthermore, it is easily seen that for all realizations given in Corollary 7 and Theorems 8 and 9, the singular value functions are uniquely determined, even though the coordinates themselves are not necessarily uniquely obtained.

Example 6. Consider the system (52) in Example 5. Let us compute the coordinate transformation $z = \Phi^{-1} \circ \bar{\Phi}(\bar{z})$ and the $\bar{\sigma}_i(\bar{z}_i)$ singular value functions according to the proof of Theorem 9. The solutions are obtained as follows:

$$\begin{aligned} z &= \Phi^{-1} \circ \bar{\Phi}(\bar{z}) = \begin{pmatrix} \frac{\sqrt{6}}{6} \bar{z}_1 + \frac{\sqrt{6}}{972} \bar{z}_1^5 \\ \frac{\sqrt{3}}{3} \bar{z}_2 + \frac{2\sqrt{3}}{243} \bar{z}_2^5 \end{pmatrix} + o(\|\bar{z}\|^5), \\ \bar{\sigma}_1(\bar{z}_1) &= 6 - \frac{2}{27} \bar{z}_1^4 + o(\|\bar{z}\|^4), \\ \bar{\sigma}_2(\bar{z}_2) &= 3 - \frac{4}{27} \bar{z}_2^4 + o(\|\bar{z}\|^4). \end{aligned}$$

This transformation converts the system (52) into the form

$$(55) \quad \begin{cases} \dot{\bar{z}} &= \bar{f}(\bar{z}) + \bar{g}(\bar{z})u, \\ y &= \bar{h}(\bar{z}) \end{cases}$$

with

$$\begin{aligned} \bar{f}(\bar{z}) &= \begin{pmatrix} -9\bar{z}_1 + \frac{7}{36}\bar{z}_1^5 + \frac{23}{9}\bar{z}_1^3\bar{z}_2^2 + \frac{31}{9}\bar{z}_1\bar{z}_2^4 \\ -9\bar{z}_2 - \frac{49}{36}\bar{z}_1^4\bar{z}_2 - \frac{17}{9}\bar{z}_1^2\bar{z}_2^3 + \frac{7}{9}\bar{z}_2^5 \end{pmatrix} + o(\|\bar{z}\|^5), \\ \bar{g}(\bar{z}) &= \begin{pmatrix} 6\sqrt{3} - \frac{19\sqrt{3}}{108}\bar{z}_1^4 - \frac{35\sqrt{3}}{27}\bar{z}_1^2\bar{z}_2^2 - \frac{5\sqrt{3}}{9}\bar{z}_2^4 & -\frac{4\sqrt{6}}{9}\bar{z}_1^3\bar{z}_2 - \frac{32\sqrt{6}}{27}\bar{z}_1\bar{z}_2^3 \\ \frac{8\sqrt{3}}{9}\bar{z}_1^3\bar{z}_2 + \frac{16\sqrt{3}}{27}\bar{z}_1\bar{z}_2^3 & 3\sqrt{6} + \frac{25\sqrt{6}}{216}\bar{z}_1^4 + \frac{25\sqrt{6}}{54}\bar{z}_1^2\bar{z}_2^2 - \frac{19\sqrt{6}}{54}\bar{z}_2^4 \end{pmatrix} \\ &\quad + o(\|\bar{z}\|^4), \\ \bar{h}(\bar{z}) &= \begin{pmatrix} 6\sqrt{3}\bar{z}_1 - \frac{19\sqrt{3}}{108}\bar{z}_1^5 - \frac{11\sqrt{3}}{27}\bar{z}_1^3\bar{z}_2^2 - \frac{7\sqrt{3}}{27}\bar{z}_1\bar{z}_2^4 \\ 3\sqrt{6}\bar{z}_2 - \frac{47\sqrt{6}}{216}\bar{z}_1^4\bar{z}_2 - \frac{31\sqrt{6}}{54}\bar{z}_1^2\bar{z}_2^3 - \frac{19\sqrt{6}}{54}\bar{z}_2^5 \end{pmatrix} + o(\|\bar{z}\|^5). \end{aligned}$$

This system has the controllability and observability functions

$$\begin{aligned} L_c(\bar{\Phi}(\bar{z})) &= \frac{1}{2} \left(\frac{\bar{z}_1^2}{\bar{\sigma}_1(\bar{z}_1)} + \frac{\bar{z}_2^2}{\bar{\sigma}_2(\bar{z}_2)} \right) + o(\|\bar{z}\|^6), \\ L_o(\bar{\Phi}(\bar{z})) &= \frac{1}{2} (\bar{z}_1^2\bar{\sigma}_1(\bar{z}_1) + \bar{z}_2^2\bar{\sigma}_2(\bar{z}_2)) + o(\|\bar{z}\|^6) \end{aligned}$$

of the form (53) and (54); that is, it is balanced in the sense of Theorem 9.

5. Model order reduction. In this section we propose balanced truncation and singular perturbation model order reduction procedures based on the balanced realizations given in the previous section, which are applicable to continuous-time and discrete-time nonlinear systems, respectively. It is shown that the proposed procedures result in reduced order models that preserve the balanced form and stability. We consider the plant systems as given in Examples 1 and 2.

5.1. Model order reduction for continuous-time systems. Consider the smooth time-invariant version of the continuous-time nonlinear system of Example 1,

$$(56) \quad \Sigma : \begin{cases} \dot{x} &= f(x, u), \\ y &= h(x, u) \end{cases}$$

with asymptotically stable equilibrium point $x = 0$ for $u = 0$, and with the Hankel operator \mathcal{H} as defined in Example 1. Suppose that Assumptions A1 and A2 hold and that we already have the coordinate transformation $z = \Phi(x)$ for one of the realizations obtained in either Theorem 7, Theorem 8, or Theorem 9. Note that all of those realizations are obtained under Assumptions A1 and A2. In the new coordinates z , the system can be described as

$$\Sigma^z : \begin{cases} \dot{z} &= f^z(z, u), \\ y &= h^z(z, u). \end{cases}$$

Here the system functions f^z and h^z , and the controllability and observability functions L_c^z and L_o^z in the coordinate z , are described by

$$(57) \quad f^z(z, u) := \left. \frac{\partial \Phi^{-1}(x)}{\partial x} \right|_{x=\Phi(z)} f(\Phi(z), u),$$

$$(58) \quad h^z(z, u) := h(\Phi(z), u),$$

$$(59) \quad L_c^z(z) := L_c(\Phi(z)),$$

$$(60) \quad L_o^z(z) := L_o(\Phi(z)).$$

The $\sigma_i(z_i)$ singular value functions are ordered as

$$(61) \quad \max_{\pm c} \sigma_i(s) > \max_{\pm c} \sigma_{i+1}(s)$$

in a neighborhood of the origin. Now let us consider the case where

$$(62) \quad \max_{\pm c} \sigma_k(s) \gg \max_{\pm c} \sigma_{k+1}(s)$$

holds for a certain $k \in \{1, \dots, n\}$. Then the state components z_1, \dots, z_k are more important in terms of the Hankel operator than z_{k+1}, \dots, z_n due to the ordering of the σ_i singular value functions; i.e., z_1, \dots, z_k cost less control energy to be reached asymptotically, and they generate more output energy than z_{k+1}, \dots, z_n .

Divide the coordinates into two parts corresponding to the division (62) as

$$(63) \quad z = (z^a, z^b) \in \mathbb{R}^n,$$

$$(64) \quad z^a := (z_1, \dots, z_k) \in \mathbb{R}^k,$$

$$(65) \quad z^b := (z_{k+1}, \dots, z_n) \in \mathbb{R}^{n-k},$$

$$(66) \quad f^z(z, u) = \begin{pmatrix} f^a(z, u) \\ f^b(z, u) \end{pmatrix}.$$

Next, divide the system Σ into two subsystems by balanced truncation (i.e., by setting either parts of the coordinates equal to zero) accordingly as follows:

$$(67) \quad \Sigma^a : \begin{cases} \dot{z}^a &= f^a(z^a, 0, u), \\ y &= h^z(z^a, 0, u), \end{cases}$$

$$(68) \quad \Sigma^b : \begin{cases} \dot{z}^b &= f^b(0, z^b, u), \\ y &= h^z(0, z^b, u). \end{cases}$$

Let \mathcal{H}^a and \mathcal{H}^b denote the Hankel operators related to the divided state-space systems (67) and (68), and let L_c^z , L_o^z , L_c^a , L_o^a , L_c^b , and L_o^b denote the controllability and observability functions of Σ in the coordinate z , and of Σ^a and Σ^b , respectively. Let σ_i , $i = 1, \dots, n$, σ_i^a , $i = 1, \dots, k$, and σ_i^b , $i = 1, \dots, n - k$, denote the singular values of the original system Σ , the reduced order systems Σ^a and Σ^b , respectively. Then we obtain the following properties, similar to the balanced truncation results for linear systems [9, 22, 32].

THEOREM 10. *Consider a continuous-time nonlinear system Σ in (56) with a Hankel operator \mathcal{H} . Suppose that Assumptions A1 and A2 hold and obtain a balanced realization on the neighborhood W as in Theorem 7 (or Theorem 8 or 9). Then the*

controllability and observability functions for the reduced order systems satisfy

$$(69) \quad L_c^a(z^a) = L_c^z(z^a, 0),$$

$$(70) \quad L_o^a(z^a) = L_o^z(z^a, 0),$$

$$(71) \quad L_c^b(z^b) = L_c^z(0, z^b),$$

$$(72) \quad L_o^b(z^b) = L_o^z(0, z^b),$$

and the singular value functions of the reduced systems satisfy

$$(73) \quad \sigma_i^a(z_i^a) = \sigma_i(z_i^a), \quad i \in \{1, 2, \dots, k\},$$

$$(74) \quad \sigma_i^b(z_i^b) = \sigma_{i+k}(z_i^b), \quad i \in \{1, 2, \dots, n-k\}.$$

The state-space systems Σ^a and Σ^b are also balanced in the sense of Theorem 7 (Theorem 8 or 9, respectively). Furthermore, there exists a neighborhood of zero $U_W \subset U$ for which the Hankel norm is preserved as

$$(75) \quad \sup_{\substack{u \in U_W \\ u \neq 0}} \frac{\|\mathcal{H}^a(u)\|}{\|u\|} = \sup_{\substack{u \in U_W \\ u \neq 0}} \frac{\|\mathcal{H}(u)\|}{\|u\|}.$$

Proof. First of all, for all realizations in Theorem 7 (Theorem 8 or 9), we have the property

$$(76) \quad z_i = 0 \iff \frac{\partial L_c(\Phi(z))}{\partial z_i} = 0 \iff \frac{\partial L_o(\Phi(z))}{\partial z_i} = 0.$$

As in (15) (see [24] for details), the observability function L_o^z of the system Σ in the coordinate z is given by a solution of a Lyapunov equation

$$\frac{\partial L_o^z(z)}{\partial z} f^z(z, 0) + \frac{1}{2} h^z(z, 0)^T h^z(z, 0) = 0.$$

Substituting $z = (z^a, 0)$ for this equation, we obtain

$$\begin{aligned} 0 &= \left(\frac{\partial L_o^z(z^a, z^b)}{\partial z^a}, \frac{\partial L_o^z(z^a, z^b)}{\partial z^b} \right) \Big|_{z^b=0} \begin{pmatrix} f^a((z^a, 0), u) \\ f^b((z^a, 0), u) \end{pmatrix} + \frac{1}{2} h^z((z^a, 0), 0)^T h^z((z^a, 0), 0) \\ &= \frac{\partial L_o^z(z^a, 0)}{\partial z^a} f^a((z^a, 0), u) + \frac{1}{2} h^z((z^a, 0), 0)^T h^z((z^a, 0), 0) \end{aligned}$$

because of (76). Clearly, this equation coincides with the Lyapunov equation for the observability function L_o^a of Σ^a . That is, we have proved the relation (70). The relation (72) can be obtained in the same way.

Next we consider the controllability function L_c^z . By the definition of \mathcal{C} and L_c in (6) and (8), it can be observed that $L_c^z(z)$ can be obtained by solving a Hamilton–Jacobi equation

$$(77) \quad \frac{\partial L_c^z(z)}{\partial z} f^z(z, u_\star(z)) + \frac{1}{2} u_\star(z)^T u_\star(z) = 0,$$

which is related to the optimal control problem in (8) (see (14) in the input-affine case). Here $u_\star(z)$ is the solution of

$$(78) \quad u = - \frac{\partial f^z(z, u)}{\partial u}^T \frac{\partial L_c^z(z)}{\partial z}^T.$$

The existence and smoothness of \mathcal{C}^\dagger in Assumption A1 implies the existence of the solution $u = u_\star(z)$ here. Substituting $z = (z^a, 0)$ for (78) yields

$$u = -\frac{\partial f^a((z^a, 0), u)}{\partial u} \frac{\partial L_c^z(z^a, 0)}{\partial z^a}^\top,$$

which is equivalent to the constraint equation for the controllability function L_c^a . Obviously, $u = u_\star(z^a, 0)$ is also the solution of this equation. Further, substituting $z = (z^a, 0)$ for (77), we obtain

$$\begin{aligned} 0 &= \left(\frac{\partial L_c^z(z^a, z^b)}{\partial z^a}, \frac{\partial L_c^z(z^a, z^b)}{\partial z^b} \right) \Big|_{z^b=0} \begin{pmatrix} f^a((z^a, 0), u_\star(z^a, 0)) \\ f^b((z^a, 0), u_\star(z^a, 0)) \end{pmatrix} + u_\star(z^a, 0)^\top u_\star(z^a, 0) \\ &= \frac{\partial L_c^z(z^a, 0)}{\partial z^a} f^z((z^a, 0), u_\star(z^a, 0)) + u_\star(z^a, 0)^\top u_\star(z^a, 0), \end{aligned}$$

which coincides with the Hamilton–Jacobi equation for the controllability function $L_c^a(z^a)$ for Σ^a . That is, we have the relation (69). The relation (71) can be obtained in the same manner.

Since the realization given in Theorem 7 (Theorem 8 or 9) is characterized only by the controllability and observability functions, the systems Σ^a and Σ^b are also balanced. Then (73)–(75) follow immediately from Theorem 7 (Theorem 8 or 9). This completes the proof. \square

Theorem 10 reveals several properties of the proposed model reduction method:

- This model reduction procedure derives *balanced* reduced order models.
- Singular value functions are preserved and, in particular, the gain of the related Hankel operator (which is called *Hankel norm*) is preserved.
- Since the controllability and observability functions are preserved, properties related to these functions, such as stability, etc. [24, 25], of the original system are preserved.

These properties are a natural nonlinear generalization of the linear case result [20], relating the Hankel theory to state-space balanced realizations and truncation. Thus, these results are far more general than the state-space balanced realization and truncation presented in [24].

Example 7. The three systems (43), (52), and (55) are all balanced in the sense of Theorems 4 and 7. Therefore, we can apply the balanced truncation procedure stated in Theorem 10 to them. As an example, it is applied to the first one, (43). Note that it also works for the other systems (52) and (55) in the same way. Since the dimension of the original system (43) is 2, the dimension of the reduced order model should be $k = 1$. Then we obtain the following 1-dimensional system:

$$\begin{cases} \dot{x}^a &= f^a(x^a) + g^a(x^a)u, \\ y &= h^a(x^a) \end{cases}$$

with

$$\begin{aligned} f^a(x^a) &= -9x^a - (x^a)^5, \\ g^a(x^a) &= (\sqrt{18 + 2(x^a)^4} \quad 0), \\ h^a(x^a) &= \begin{pmatrix} \frac{(6x^a - 2(x^a)^5)\sqrt{18 + 2(x^a)^4}}{1 + (x^a)^4} \\ 0 \end{pmatrix}. \end{aligned}$$

The controllability, observability, and the Hankel singular value functions for this system can be computed as follows:

$$\begin{aligned} L_c^a(x^a) &= \frac{1}{2}(x^a)^2, \\ L_o^a(x^a) &= \frac{1}{2}(x^a)^2 \tau_1^a(x^a) = \frac{1}{2}(x^a \sigma_1^a(x^a))^2, \\ \sigma_1^a(x^a) &= \sigma_1(x^a) = 2\sqrt{\frac{9 + (x^a)^4}{1 + (x^a)^4}}, \end{aligned}$$

which confirms the outcome of Theorem 10.

5.2. Model order reduction for discrete-time systems. This section proposes model order reduction based on the balanced representation of section 4 for discrete-time nonlinear systems. Consider the time-invariant version of the discrete-time nonlinear system of Example 2,

$$(79) \quad \Sigma : \begin{cases} x(t+1) &= f(x(t), u(t)), \\ y(t) &= h(x(t), u(t)) \end{cases}$$

with asymptotically stable equilibrium point $x = 0$ for $u = 0$, and with the Hankel operator \mathcal{H} as defined in Example 2. As in the example, we suppose that the $x(t+1) = f(x(t), u(t))$ is invertible with respect to $x(t)$; that is, there exists a function f^{-1} satisfying

$$f(f^{-1}(x, u), u) = x \quad \forall x \in \mathbb{R}^n, u \in \mathbb{R}^m.$$

As a preparation for the model order reduction for discrete-time systems, we need to characterize the observability and controllability functions $L_o(x)$ and $L_c(x)$ by algebraic equations which are similar to the Hamilton–Jacobi equations in the continuous-time case. We introduce a modified version of Lemma 4.1 and Theorem 4.4 in [17].

LEMMA 2. Suppose that $x = 0$ of the system

$$x(t+1) = f(x(t), 0)$$

is asymptotically stable. Then the observability function $L_o(x)$ in (9) exists if and only if

$$(80) \quad \check{L}_o(f(x, 0)) - \check{L}_o(x) + \frac{1}{2}h(x, 0)^T h(x, 0) = 0, \quad \check{L}_o(0) = 0$$

has a solution $\check{L}_o(x)$. If it exists, then $L_o(x) = \check{L}_o(x)$ holds.

Proof. Necessity is proved first. Suppose that the observability function $L_o(x)$ exists. Then the definition of the observability function (9) implies that

$$\begin{aligned} L_o(x(0)) &= \frac{1}{2} \sum_{t=0}^{\infty} h(x(t), 0)^T h(x(t), 0) \\ &= \frac{1}{2} \sum_{t=1}^{\infty} h(x(t), 0)^T h(x(t), 0) + \frac{1}{2} h(x(0), 0)^T h(x(0), 0) \\ &= L_o(x(1)) + \frac{1}{2} h(x(0), 0)^T h(x(0), 0) \\ &= L_o(f(x(0), 0)) + \frac{1}{2} h(x(0), 0)^T h(x(0), 0). \end{aligned}$$

This equation has to hold for an arbitrary initial state $x(0)$, that is, it satisfies (80) with $\check{L}_o(x) = L_o(x)$ since $L_o(0) = 0$. This proves the necessity.

Next, sufficiency is proved. Suppose that (80) has a smooth solution $\check{L}_o(x)$. Using the notation $F(x) := f(x, 0)$, (80) implies that

$$\begin{aligned}\check{L}_o(x) &= \check{L}_o(F(x)) + \frac{1}{2}h(x, 0)^T h(x, 0) \\ &= \check{L}_o(F(F(x))) + \frac{1}{2}h(x, 0)^T h(x, 0) + \frac{1}{2}h(F(x), 0)^T h(F(x), 0) \\ &= \lim_{k \rightarrow \infty} \left(\check{L}_o(F^k(x)) + \frac{1}{2} \sum_{i=0}^k h(F^i(x), 0)^T h(F^i(x), 0) \right) \\ &= \lim_{k \rightarrow \infty} \check{L}_o(F^k(x)) + L_o(x) \\ &= L_o(x).\end{aligned}$$

The last equation holds because the system $x(t+1) = F(x(t))$ is asymptotically stable and because $\check{L}_o(0) = 0$. This completes the proof. \square

This result is a natural nonlinear generalization of the linear case result. In the linear case, the dynamics (25) is given by

$$(81) \quad \Sigma : \begin{cases} x(t+1) &= Ax(t) + Bu(t), \\ y(t) &= Cx(t) + Du(t) \end{cases}$$

with A , B , C , and D of appropriate size. Then the observability function is quadratic, i.e.,

$$L_o(x) = \frac{1}{2}x^T Q_d x.$$

Equation (80) reduces to the Lyapunov equation

$$A^T Q_d A - Q_d + C^T C = 0,$$

where Q_d is the observability Gramian of the linear discrete-time system.

A similar result for the controllability function is obtained as follows. Let us consider an optimal control problem minimizing a cost function

$$(82) \quad L_c(x^0) = \min_{\substack{v \in \ell_2(\mathbb{Z}) \\ x(-\infty)=0, \quad x(0)=x^0}} \frac{1}{2} \sum_{t=0}^{\infty} \|v(t)\|^2$$

for the dynamics of \mathcal{C} ,

$$x(t+1) = f(x(t), u(t)), \quad u(t) = v(-t-1), \quad t = -1, -2, \dots$$

Let us denote the input u achieving the minimization in (82) by $u(t) = u^*(x(t+1))$, which depends on $x(t+1)$ since this is an optimal control problem with respect to the reverse time. Then the dynamics of $\mathcal{C}^\dagger : x^0 \mapsto v$ becomes

$$\mathcal{C}^\dagger : \begin{cases} x(t+1) &= f(x(t), u^*(x(t+1))), \quad x(0) = x^0, \quad t = -1, -2, \dots, \\ v(-t-1) &= u^*(x(t+1)), \end{cases}$$

which reduces to

$$\mathcal{C}^\dagger : \begin{cases} \bar{x}(\bar{t}+1) &= f^{-1}(\bar{x}(\bar{t}), u^*(\bar{x}(\bar{t}))), \quad \bar{x}(0) = x^0, \quad \bar{t} = 0, 1, 2, \dots, \\ v(\bar{t}) &= u^*(\bar{x}(\bar{t})) \end{cases}$$

with $\bar{t} = -t$ and $\bar{x}(\bar{t}) = x(t)$. Now a modified version of Lemma 4.2 and Theorem 4.5 in [17] is given as follows.

LEMMA 3. *Suppose that $x = 0$ of the feedback system*

$$x(t+1) = f^{-1}(x(t), u^*(x(t)))$$

is asymptotically stable. Then a controllability function $L_c(x)$ in (8) exists if and only if

$$(83) \quad \check{L}_c(f^{-1}(x, u^*(x))) - \check{L}_c(x) + \frac{1}{2}u^*(x)^T u^*(x) = 0, \quad \check{L}_c(0) = 0,$$

has a solution $\check{L}_c(x)$. If it exists, then $L_c(x) = \check{L}_c(x)$ holds.

Proof. This lemma can be proved as a corollary of Lemma 2 by substituting \mathcal{C}^\dagger for \mathcal{O} . \square

These results are also a natural generalization of the continuous-time Lyapunov/Hamilton–Jacobi equations (80) and (83) that characterize the observability and controllability functions.

The characterization of the discrete-time observability and controllability functions in the above two lemmas is useful for model order reduction. However, performing balanced truncation as in the continuous-time case will not result in reduced order systems that preserve the balanced realization properties. A way to circumvent this is to consider singular perturbation model reduction based on the balanced representation, similar to the linear case; see, e.g., [13, 15]. As in the previous section, let us now suppose that Assumptions A1 and A2 hold and that we already have the coordinate transformation $z = \Phi(x)$ in the neighborhood W for one of the realizations obtained in Theorems 8 and 9.

Consider the system Σ in (25) and suppose that the system is balanced in the sense of Theorem 8 or 9. The original dynamics can be described in the z coordinates as follows:

$$(84) \quad \Sigma : \begin{cases} z(t+1) &= f^z(z(t), u(t)), \\ y(t) &= h^z(z(t), u(t)). \end{cases}$$

Here the system functions f^z and h^z , and the controllability and observability functions L_c^z and L_o^z in the coordinate z , are described by (58)–(60) and

$$f^z(z, u) := \Phi^{-1} \circ f(\Phi(z), u).$$

As in the continuous-time case, suppose that the singular value functions are ordered as in (61) and (62) and divide the state-space as in (63)–(66). Then, accordingly, we obtain two reduced order systems by a singular perturbation method:

$$\begin{aligned} \Sigma^a : & \begin{cases} z^a(t+1) &= f^a(z^a(t), z^b(t), u(t)), \\ z^b(t) &= f^b(z^a(t), z^b(t), u(t)), \\ y(t) &= h(z^a(t), z^b(t), u(t)), \end{cases} \\ \Sigma^b : & \begin{cases} z^a(t) &= f^a(z^a(t), z^b(t), u(t)), \\ z^b(t+1) &= f^b(z^a(t), z^b(t), u(t)), \\ y^b(t) &= h(z^a(t), z^b(t), u(t)). \end{cases} \end{aligned}$$

Here we suppose that

$$(85) \quad z^a = f^a(z^a, z^b, u)$$

has a unique solution

$$(86) \quad z^a = \hat{f}^a(z^b, u),$$

which describes the stationary state of the subsystem (85) for a given input (z^b, u) . We also assume that the equation

$$(87) \quad z^b = f^b(z^a, z^b, u)$$

has a unique solution

$$(88) \quad z^b = \hat{f}^b(z^a, u)$$

describing the stationary state of the subsystem (87) for a given input (z^a, u) . Note that simple sufficient conditions for the existence of the functions \hat{f}^a and \hat{f}^b are

$$\det \left(I - \frac{\partial f^a}{\partial z^a} \right) \neq 0,$$

$$\det \left(I - \frac{\partial f^b}{\partial z^b} \right) \neq 0,$$

respectively. Then we obtain explicit forms

$$(89) \quad \Sigma^a : \begin{cases} z^a(t+1) &= \bar{f}^a(z^a(t), u(t)), \\ y(t) &= \bar{h}^a(z^a(t), u(t)), \end{cases}$$

$$(90) \quad \Sigma^b : \begin{cases} z^b(t+1) &= \bar{f}^b(z^b(t), u(t)), \\ y^b(t) &= \bar{h}^b(z^b(t), u(t)) \end{cases}$$

with

$$\begin{aligned} \bar{f}^a(z^a(t), u(t)) &:= f^a(z^a(t), \hat{f}^b(z^a(t), u(t)), u(t)), \\ \bar{h}^a(z^a(t), u(t)) &:= h(z^a(t), \hat{f}^b(z^a(t), u(t)), u(t)), \\ \bar{f}^b(z^b(t), u(t)) &:= f^b(\hat{f}^a(z^b(t), u(t)), z^b(t), u(t)), \\ \bar{h}^b(z^b(t), u(t)) &:= h(\hat{f}^a(z^b(t), u(t)), z^b(t), u(t)) \end{aligned}$$

by substituting (86) and (88) for Σ in (84). For these reduced order systems, we can prove the following properties.

THEOREM 11. *Consider a discrete-time nonlinear system Σ in (79) with a Hankel operator \mathcal{H} . Suppose that Assumptions A1 and A2 hold and obtain a balanced realization of Theorem 8 (or Theorem 9) in a neighborhood W . Then the controllability and observability functions for the reduced systems satisfy*

$$\begin{aligned} L_c^a(z^a) &= L_c^z(z^a, 0), \\ L_o^a(z^a) &= L_o^z(z^a, 0), \\ L_c^b(z^b) &= L_c^z(0, z^b), \\ L_o^b(z^b) &= L_o^z(0, z^b), \end{aligned}$$

and the singular value functions of the reduced systems satisfy

$$\begin{aligned} \sigma_i^a(z_i^a) &= \sigma_i(z_i^a), \quad i \in \{1, 2, \dots, k\}, \\ \sigma_i^b(z_i^b) &= \sigma_{i+k}(z_i^b), \quad i \in \{1, 2, \dots, n-k\}. \end{aligned}$$

The state-space systems Σ^a and Σ^b are also balanced in the sense of Theorem 8 (or Theorem 9, respectively). Furthermore, there exists a neighborhood of zero $U_W \subset U$ for which the Hankel norm is preserved as

$$\sup_{\substack{u \in U_W \\ u \neq 0}} \frac{\|\mathcal{H}^a(u)\|}{\|u\|} = \sup_{\substack{u \in U_W \\ u \neq 0}} \frac{\|\mathcal{H}(u)\|}{\|u\|}.$$

Proof. First of all, since the system is balanced in the sense of Theorem 8 or 9, the controllability and observability functions can be separated as

$$\begin{aligned} L_c^z(z) &= L_c(\Phi(z)) = \sum_{i=1}^n l_i^c(z_i), \\ L_o^z(z) &= L_o(\Phi(z)) = \sum_{i=1}^n l_i^o(z_i) \end{aligned}$$

with scalar functions $l_i^c(z_i)$ and $l_i^o(z_i)$, $i = 1, \dots, n$. Then it implies that $L_o^z(z)$ can be divided into two parts

$$(91) \quad L_o^z(z) = \check{L}_o^a(z^a) + \check{L}_o^b(z^b)$$

with $\check{L}_o^a(z^a) = L_o^z(z^a, 0)$ and $\check{L}_o^b(z^b) = L_o^z(0, z^b)$. On the other hand, (85)–(88) imply that

$$(92) \quad f^a(\hat{f}^a(z^b, u), z^b, u) = \hat{f}^a(z^b, u),$$

$$(93) \quad f^b(z^a, \hat{f}^b(z^a, u), u) = \hat{f}^b(z^a, u).$$

Let us substitute (88) for (80) in Lemma 2. Then we obtain

$$\begin{aligned} 0 &= \left[L_o^z(f(z, 0)) - L_o^z(z) + \frac{1}{2} h(z, 0)^T h(z, 0) \right] \Big|_{z^b = \hat{f}^b(z^a, u)} \\ &= L_o^z(f(z^a, \hat{f}^b(z^a, 0), 0)) - L_o^z(z^a, \hat{f}^b(z^a, 0)) + \frac{1}{2} \|h(z^a, \hat{f}^b(z^a, 0), 0)\|^2 \\ &= \left(\check{L}_o^a(f^a(z^a, \hat{f}^b(z^a, 0), 0)) + \check{L}_o^b(f^b(z^a, \hat{f}^b(z^a, 0), 0)) \right) - \left(\check{L}_o^a(z^a) + \check{L}_o^b(\hat{f}^b(z^a, 0)) \right) \\ &\quad + \frac{1}{2} \|h(z^a, \hat{f}^b(z^a, 0), 0)\|^2 \\ &= \check{L}_o^a(\bar{f}^a(z^a, 0)) - \check{L}_o^a(z^a) + \frac{1}{2} \|\bar{h}^a(z^a, 0)\|^2. \end{aligned}$$

Here the third equation follows from (91), and the last equation follows from (92) and (93). Then Lemma 2 implies that $\check{L}_o^a(z^a) = L_o^z(z^a, 0)$ is the observability function of the system Σ^a . Further, it is easily seen that $\check{L}_o^b(z^b) = L_o^z(0, z^b)$ is the observability function of Σ^b by substituting (86).

In a similar way, as in the proof of Lemma 3, by identifying \mathcal{C}^\dagger with \mathcal{O} , we can prove that the controllability function is divided into

$$(94) \quad L_c^z(z) = \check{L}_c^a(z^a) + \check{L}_c^b(z^b),$$

which proves the former part of the theorem. The latter part follows as in the proof of Theorem 10. This completes the proof. \square

This theorem is a discrete-time counterpart of the continuous-time result in Theorem 10, although we use a singular perturbation reduction procedure. It is proved that this model order reduction procedure preserves the controllability and observability functions and their properties.

6. Conclusion. In this paper, singular value analysis of Hankel operators for both continuous-time and discrete-time finite and infinite dimensional nonlinear systems has been discussed. Singular value analysis of operators clarifies the gain structure of a given operator. Here it is proved that this structure of smooth Hankel operators of general nonlinear systems can be characterized by a simple equation in terms of the state. This result can be utilized for balanced realization and model order reduction for finite dimensional continuous-time and discrete-time input-nonaffine nonlinear systems. Furthermore, we have derived a precise balanced realization for nonlinear systems, whereas the existing approach gave only an input-normal realization. Moreover, based on the proposed balanced realization for general nonlinear systems, model order reduction procedures for both continuous-time and discrete-time systems are derived. In these methods, several important properties of the original system, such as stability, controllability, observability, and the gain property, are preserved.

Appendix.

Proof of Lemma 1.

Proof. First, the system is brought into the form of Theorem 7 on the coordinate x , that is,

$$(95) \quad L_c(x) = \frac{1}{2} (x_1^2 + x_2^2),$$

$$(96) \quad L_o(x) = \frac{1}{2} (x_1^2 \tau_1(x) + x_2^2 \tau_2(x)),$$

$$(97) \quad \sigma_1(x_1)^2 = \tau_1(x_1, 0),$$

$$(98) \quad \sigma_2(x_2)^2 = \tau_2(0, x_2),$$

$$(99) \quad x_i = 0 \iff \frac{\partial L_o}{\partial x_i} = 0.$$

Let $\tilde{L}_o(z)$ denote the balanced observability function, that is,

$$\tilde{L}_o(z) := \frac{1}{2} ((z_1 \sigma_1(z_1))^2 + (z_2 \sigma_2(z_2))^2).$$

What we have to prove is the existence of a coordinate transformation $x = \Phi(z)$ converting $L_c(x)$ and $L_o(x)$ in the above equations into $L_c(z)$ and $\tilde{L}_o(z)$. Hence the coordinate transformation $x = \Phi(z)$ is a solution of the equations

$$(100) \quad F_c(x, z) := L_c(x) - L_c(z) = 0,$$

$$(101) \quad F_o(x, z) := L_o(x) - \tilde{L}_o(z) = 0.$$

Now define the polar coordinates

$$\theta := \begin{pmatrix} r \\ \theta_1 \end{pmatrix} = \begin{pmatrix} \sqrt{x_1^2 + x_2^2} \\ \text{atan2}(x_2, x_1) \end{pmatrix} =: \Theta(x),$$

$$\varphi := \begin{pmatrix} s \\ \varphi_1 \end{pmatrix} = \begin{pmatrix} \sqrt{z_1^2 + z_2^2} \\ \text{atan2}(z_2, z_1) \end{pmatrix} =: \Theta(z).$$

Then (100) and (101) can be converted into the polar coordinates as follows:

$$(102) \quad F_c(\Theta^{-1}(\theta), \Theta^{-1}(\varphi)) = 0,$$

$$(103) \quad F_o(\Theta^{-1}(\theta), \Theta^{-1}(\varphi)) = 0.$$

If we can find a smooth solution $\theta = \Psi(\varphi)$, then the coordinate transformation $x = \Phi(z)$ can be obtained by $x = \Phi(z) = \Theta^{-1} \circ \Psi \circ \Theta(z)$. In what follows, we will prove the existence of a smooth coordinate transformation $\theta = \Psi(\varphi)$. Note that (102) is satisfied if and only if $r = s$, that is,

$$F_c(\Theta^{-1}(s, \theta_1), \Theta^{-1}(s, \varphi_1)) \equiv 0$$

holds. Hence, what we have to solve is (103), namely, we need to find a solution $\theta_1 = \psi(s, \varphi_1)$ satisfying

$$(104) \quad F_o(\Theta^{-1}(s, \theta_1), \Theta^{-1}(s, \varphi_1)) = 0.$$

Then the function Ψ is obtained by

$$(105) \quad \begin{pmatrix} r \\ \theta_1 \end{pmatrix} = \Psi(\varphi) := \begin{pmatrix} s \\ \psi(s, \varphi_1) \end{pmatrix}.$$

Here we will prove the existence and invertibility of a scalar function $\theta_1 = \psi(s, \varphi_1)$ for any fixed (small enough) s . The derivative of F_o in (104) with respect to θ_1 and φ_1 can be calculated as

$$\begin{aligned} \frac{\partial F_o}{\partial \theta_1} &= \frac{\partial F_o(x, z)}{\partial x} \frac{\partial \Theta^{-1}(\theta)}{\partial \theta_1} = \frac{\partial L_o(x)}{\partial x} \frac{\partial \Theta^{-1}(\theta)}{\partial \theta_1} \\ &= -\frac{\partial L_o(x)}{\partial x_1} s \sin \theta_1 + \frac{\partial L_o(x)}{\partial x_2} s \cos \theta_1 \\ (106) \quad &= -\frac{\partial L_o(x)}{\partial x_1} x_2 + \frac{\partial L_o(x)}{\partial x_2} x_1, \end{aligned}$$

$$(107) \quad \frac{\partial F_o}{\partial \varphi_1} = \frac{\partial \tilde{L}_o(z)}{\partial z_1} z_2 - \frac{\partial \tilde{L}_o(z)}{\partial z_2} z_1.$$

The relationship (99) and Lemma 2.1 in [18] imply that there exist smooth scalar functions $\ell_i(x)$ and $\tilde{\ell}_i(z_i)$ satisfying

$$\begin{aligned} \frac{\partial L_o(x)}{\partial x_i} &= x_i \ell_i(x), \\ \frac{\partial \tilde{L}_o(z)}{\partial z_i} &= z_i \tilde{\ell}_i(z_i), \end{aligned}$$

which reduce (106) and (107) to

$$(108) \quad \frac{\partial F_o}{\partial \theta_1} = -x_1 x_2 (\ell_1(x) - \ell_2(x)),$$

$$(109) \quad \frac{\partial F_o}{\partial \varphi_1} = z_1 z_2 (\tilde{\ell}_1(z_1) - \tilde{\ell}_2(z_2)).$$

The functions ℓ_i and $\tilde{\ell}_i$ coincide at the origin with the Hankel singular value σ_i of the Jacobian linearization of the system, i.e.,

$$\ell_i(0) = \tilde{\ell}_i(0) = \sigma_i^2.$$

Assumption A2 guarantees that there exists a neighborhood of the origin, where $\ell_1(x) > \ell_2(x)$, $\tilde{\ell}_1(z_1) > \tilde{\ell}_2(z_2)$ hold. Hence (108) and (109) imply that

$$(110) \quad \frac{\partial F_o}{\partial \theta_1} = 0 \iff x_1 x_2 = 0 \iff \theta_1 = 0 \bmod \frac{\pi}{2},$$

$$(111) \quad \frac{\partial F_o}{\partial \varphi_1} = 0 \iff z_1 z_2 = 0 \iff \varphi_1 = 0 \bmod \frac{\pi}{2},$$

hold in the neighborhood of the origin. On the other hand, (96), (97), and (98) imply

$$\theta_1 = 0 \bmod \frac{\pi}{2} \Rightarrow L_o(x) = \tilde{L}_o(x) \Rightarrow F_o(x, x) = 0.$$

That is, the coordinate transformation $x = \Phi(z)$ (and $\theta = \Psi(\varphi)$ also) has to coincide with the identity on the axes $x = (x_1, 0)$ and $x = (0, x_2)$. Let us consider the map

$$\theta_1 = \psi(s, \varphi_1)$$

on a region $0 \leq \varphi_1 \leq \pi/2$. Then $\varphi_1 = 0 \Rightarrow \theta_1 = 0$ and $\varphi_1 = \pi/2 \Rightarrow \theta_1 = \pi/2$ hold. The intermediate value theorem with the above properties implies that, for any $\varphi_1 \in (0, \pi/2)$ and any $s \neq 0$, there exists a corresponding $\theta_1 \in (0, \pi/2)$ and vice versa. Furthermore, the implicit function theorem and the relationships (110) and (111) imply that the mapping $\varphi_1 \mapsto \theta_1$ is a diffeomorphism at least for all $\varphi_1 \in (0, \pi/2)$. Hence what remains to be proved is the smoothness of ψ at the points where $\varphi_1 = 0 \bmod \pi/2$.

The determinant of the Jacobian matrix of Ψ defined by (105) is given by

$$\det \frac{\partial \Psi(\varphi)}{\partial \varphi} = \det \begin{pmatrix} 1 & 0 \\ \frac{\partial \psi}{\partial s} & \frac{\partial \psi}{\partial \varphi_1} \end{pmatrix} = \frac{\partial \psi(s, \varphi_1)}{\partial \varphi_1}.$$

Therefore the invertibility of the mapping Ψ is implied by proving $\partial \psi / \partial \varphi_1$ (or $\partial \theta_1 / \partial \varphi_1$) is not zero. Here the implicit function theorem implies

$$\frac{\partial \theta_1}{\partial \varphi_1} = - \frac{\frac{\partial F_o}{\partial \varphi_1}}{\frac{\partial F_o}{\partial \theta_1}} = \frac{z_1 z_2 (\tilde{\ell}_1(z_1) - \tilde{\ell}_2(z_2))}{x_1 x_2 (\ell_1(x) - \ell_2(x))} = \frac{\sin \varphi_1 \cos \varphi_1 (\tilde{\ell}_1(z_1) - \tilde{\ell}_2(z_2))}{\sin \theta_1 \cos \theta_1 (\ell_1(x) - \ell_2(x))}.$$

But this is indefinite at $\varphi_1 = 0$ ($\theta_1 = 0$). Using l'Hospital's theorem, we can obtain

$$\begin{aligned} \lim_{\varphi_1 \rightarrow 0} \frac{\partial \theta_1}{\partial \varphi_1} &= - \lim_{\varphi_1 \rightarrow 0} \frac{\frac{\partial F_o}{\partial \varphi_1}}{\frac{\partial F_o}{\partial \theta_1}} = - \lim_{\varphi_1 \rightarrow 0} \frac{\frac{\partial^2 F_o}{\partial \varphi_1^2}}{\frac{\partial^2 F_o}{\partial \varphi_1 \partial \theta_1}} \\ &= - \lim_{\varphi_1 \rightarrow 0} \frac{\frac{\partial^2 F_o}{\partial \varphi_1^2}}{\frac{\partial^2 F_o}{\partial \theta_1^2} \frac{\partial \theta_1}{\partial \varphi_1}} = - \lim_{\varphi_1 \rightarrow 0} \frac{\frac{\partial^2 F_o}{\partial \varphi_1^2}}{\frac{\partial^2 F_o}{\partial \theta_1^2}} \lim_{\varphi_1 \rightarrow 0} \frac{1}{\frac{\partial \theta_1}{\partial \varphi_1}}. \end{aligned}$$

Therefore we have

$$\lim_{\varphi_1 \rightarrow 0} \frac{\partial \theta_1}{\partial \varphi_1} = \left(- \lim_{\varphi_1 \rightarrow 0} \frac{\frac{\partial^2 F_o}{\partial \varphi_1^2}}{\frac{\partial^2 F_o}{\partial \theta_1^2}} \right)^{\frac{1}{2}} = \sqrt{\frac{\tilde{\ell}_1(s) - \tilde{\ell}_2(0)}{\ell_1(s, 0) - \ell_2(s, 0)}}$$

since $\partial\theta_1/\partial\varphi_1$ is nonnegative. This limit exists and takes a positive value for small enough s since $\tilde{\ell}_i(0) = \ell_i(0, 0) = \sigma_i^2$. Therefore the mapping Ψ is invertible and differentiable. Higher order derivatives can be derived easily since the functions L_c and L_o are smooth, which suggests that $\theta = \Psi(\varphi)$ is also smooth. Similar relationships hold in the other cases where $\theta_1 = \varphi_1 = \pm(\pi/2), \pi$. This completes the proof. \square

Proof of Theorem 8.

Proof. As in the proof of Lemma 1, it is assumed without loss of generality that the system is already balanced in the sense of Theorem 7 on the coordinate x . The theorem is proved by induction with respect to the dimension n .

(i) Case $n = 1$ holds obviously.

(ii) Case $n = 2$ is proved in Lemma 1.

(iii) Case $n = k$: Suppose that the theorem holds in the case $n = k - 1$. Let us define truncated vectors $(\cdot)_i$ and $(\check{\cdot})_i$ for a given vector $x = (x_1, x_2, \dots, x_k) \in \mathbb{R}^k$ by

$$\begin{aligned}\check{x}_i &:= (x_1, \dots, x_{i-1}, 0, x_{i+1}, \dots, x_k) \in \mathbb{R}^k, \\ \check{x}_i &:= (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_k) \in \mathbb{R}^{k-1}.\end{aligned}$$

First of all, let us apply the theorem to the system restricted to the subspace $\{x \mid x_k = 0\}$. The theorem in the case $n = k - 1$ (assumed above) implies that there exists a coordinate transformation $\check{x}_k = \Psi_k(\check{z}_k)$ satisfying

$$\begin{aligned}L_c(\Psi_k(\check{z}_k), 0) &= \frac{1}{2} \check{z}_k^T \check{z}_k, \\ L_o(\Psi_k(\check{z}_k), 0) &= \frac{1}{2} \sum_{i=1}^{k-1} (z_i \sigma_i(z_i))^2.\end{aligned}$$

As in the proof of Lemma 1, in order to construct a coordinate transformation preserving the input-normal form, let us define the generalized polar coordinate

$$\theta := \begin{pmatrix} r \\ \theta_1 \\ \vdots \\ \theta_{k-1} \end{pmatrix} = \begin{pmatrix} (x_1^2 + x_2^2 + \dots + x_n^2)^{1/2} \\ \text{atan2}(x_2, x_1) \\ \vdots \\ \text{atan2}(x_k, (x_1^2 + \dots + x_{k-1}^2)^{1/2}) \end{pmatrix} =: \Theta(x).$$

By definition, $x_1 = 0 \Leftrightarrow \theta_1 = \pi/2$ and $x_{i+1} = 0 \Leftrightarrow \theta_i = 0$. We also define the generalized polar coordinate corresponding to z by $\varphi := (s, \varphi_1, \dots, \varphi_{k-1}) := \Theta(z)$. Then the function Ψ_k has to satisfy

$$(112) \quad \check{\theta}_k = \check{\Theta}_k \circ \Psi_k \circ \check{\Theta}_k^{-1}(\check{\varphi}_k) =: \check{\Psi}_k(\check{\varphi}_k).$$

In these coordinates, consider a rotational matrix $R(\varphi, \theta) \in \mathbb{R}^{k \times k}$ changing the polar coordinate θ into φ with $s = r$ defined by $R(\varphi, \theta) := R_{k-1}(\varphi_{k-1}) \cdots R_1(\varphi_1) R_1(-\theta_1) \cdots R_{k-1}(-\theta_{k-1})$ with $R_i(\theta_i)$ being the rotation matrices for the component angles θ_i ,

$i = 1, \dots, k-1$, defined by

$$R_1(\theta_1) := \begin{pmatrix} 1 & 0 & 0 & & 0 \\ 0 & \cos \theta_1 & -\sin \theta_1 & & \\ 0 & \sin \theta_1 & \cos \theta_1 & & \\ & & & 1 & 0 \\ & & & & \ddots \\ 0 & & & 0 & 1 \end{pmatrix} \in \mathbb{R}^{k \times k},$$

$$R_2(\theta_2) := \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \cos \theta_2 & 0 & -\sin \theta_2 & \\ 0 & 0 & 1 & 0 & \\ 0 & \sin \theta_2 & 0 & \cos \theta_2 & \\ & & & 1 & 0 \\ & & & & \ddots \\ 0 & & & 0 & 1 \end{pmatrix} \in \mathbb{R}^{k \times k}.$$

$$\vdots$$

Using the mapping $\check{\varphi}_k = \tilde{\Psi}_k^{-1}(\check{\theta}_k)$ defined on S^{k-1} , we can construct a coordinate transformation on S^k by

$$\begin{aligned} \varphi &= \Psi_k^{-1}(\theta) \\ &:= (\lambda(\theta_k) \tilde{\Psi}_k^{-1}(\check{\theta}_k) + (1 - \lambda(\theta_k)) \check{\theta}_k, \theta_k), \end{aligned}$$

where λ is a smooth scalar function with an appropriate constant ϵ ($0 < \epsilon < \pi/2$) (cf. [19]):

$$\lambda(s) := \begin{cases} 0 & (s \geq 2\epsilon), \\ \frac{\exp(\epsilon/s)}{\exp(-\epsilon/(s+\epsilon)) + \exp(\epsilon/s)} & (\epsilon \leq s \leq 2\epsilon), \\ 1 & (-\epsilon \leq s \leq \epsilon), \\ \frac{\exp(-\epsilon/(s+2\epsilon))}{\exp(-\epsilon/(s+2\epsilon)) + \exp(\epsilon/(s+\epsilon))} & (-2\epsilon \leq s \leq -\epsilon), \\ 0 & (s \leq -2\epsilon). \end{cases}$$

It is readily observed that $\Psi_k(\check{\varphi}_k) = \tilde{\Psi}_k(\check{\varphi}_k)$. Furthermore, a coordinate transformation on \mathbb{R}^k can be constructed by

$$x = \Phi_k(\xi) := R(\Psi_k(\varphi), \varphi)\xi = R(\Psi_k \circ \Theta(\xi), \Theta(\xi))\xi,$$

which is defined in a neighborhood of the origin. By its construction, this coordinate transformation $x = \Phi_k(\xi)$ satisfies

$$(113) \quad L_o(\Phi_k(\check{\xi}_k)) = \frac{1}{2} \sum_{i=1}^{k-1} (\xi_i \sigma_i(\xi_i))^2$$

without losing the properties achieved in Theorem 7.

Next let us construct a coordinate transformation $\xi = \Phi_{k-1}(\zeta)$, which achieves the balanced realization in the subspace $\{\xi \mid \xi_{k-1} = 0\}$, that is,

$$(114) \quad L_o(\Phi_k \circ \Phi_{k-1}(\check{\zeta}_{k-1})) = \frac{1}{2} \sum_{\substack{i=1 \\ i \neq k-1}}^k (\zeta_i \sigma_i(\zeta_i))^2.$$

Since the subspace $\{\xi \mid \xi_{k-1} = \xi_k = 0\}$ is already balanced in the sense that (113) already holds, Φ_k can be chosen in such a way that it coincides with the identity on $\{\zeta \mid \zeta_{k-1} = \zeta_k = 0\}$. This fact reveals that the following property also holds:

$$(115) \quad L_o(\Phi_k \circ \Phi_{k-1}(\check{\zeta}_k)) = \frac{1}{2} \sum_{i=1}^{k-1} (\zeta_i \sigma_i(\zeta_i))^2.$$

Furthermore, since the coordinate transformations constructed here preserve the properties in Theorem 7, we have

$$(116) \quad L_c(\Phi_k \circ \Phi_{k-1}(\zeta)) = \frac{1}{2} \zeta^T \zeta,$$

$$(117) \quad L_o(\Phi_k \circ \Phi_{k-1}(\zeta)) = \frac{1}{2} \sum_{i=1}^k (\zeta_i \bar{\tau}_i(\zeta))^2,$$

$$(118) \quad \sigma_i(\zeta_i)^2 = \bar{\tau}_i(0, \dots, 0, \underbrace{\zeta_i}_{i\text{th}}, 0, \dots, 0),$$

$$(119) \quad \zeta_i = 0 \iff \frac{\partial L_o(\Phi_k \circ \Phi_{k-1}(\zeta))}{\partial \zeta_i} = 0.$$

Now let us define virtual controllability and observability functions of ζ_{k-1} and ζ_k by regarding the other variables ζ_i ($i = 1, 2, \dots, k-2$) as constants:

$$\bar{L}_c(\zeta_{k-1}, \zeta_k) := \frac{1}{2} (\zeta_{k-1}^2 + \zeta_k^2),$$

$$\bar{L}_o(\zeta_{k-1}, \zeta_k) := L_o(\Phi_k \circ \Phi_{k-1}(\zeta)) - \frac{1}{2} \sum_{i=1}^{k-2} (\zeta_i \sigma_i(\zeta_i))^2.$$

Note that, due to the relationships (114), (115), and (119), this function satisfies the following properties, at least in a neighborhood of the origin for any ζ_i 's ($i = 1, 2, \dots, k-2$):

$$\begin{aligned} \bar{L}_o(\zeta_{k-1}, \zeta_k) &\geq 0, \\ \bar{L}_o(\zeta_{k-1}, \zeta_k) = 0 &\iff \zeta_{k-1} = \zeta_k = 0. \end{aligned}$$

The properties (116)–(119) imply that these functions are already balanced in the sense of Theorem 7. Therefore, application of Lemma 1 to this pair of functions on the state space (ζ_{k-1}, ζ_k) proves the existence of a coordinate transformation $(\zeta_{k-1}, \zeta_k) = \bar{\phi}(\bar{\zeta}_{k-1}, \bar{\zeta}_k)$ (which also depends on ζ_i ($i = 1, 2, \dots, k-2$)) satisfying

$$\begin{aligned} \bar{L}_c(\bar{\phi}(\bar{\zeta}_{k-1}, \bar{\zeta}_k)) &= \frac{1}{2} (\bar{\zeta}_{k-1}^2 + \bar{\zeta}_k^2), \\ \bar{L}_o(\bar{\phi}(\bar{\zeta}_{k-1}, \bar{\zeta}_k)) &= \frac{1}{2} ((\bar{\zeta}_{k-1} \sigma_{k-1}(\bar{\zeta}_{k-1}))^2 + (\bar{\zeta}_k \sigma_k(\bar{\zeta}_k))^2). \end{aligned}$$

Let us define a coordinate transformation on \mathbb{R}^k by

$$\begin{aligned} x &= \Phi(z) := \Phi_k \circ \Phi_{k-1} \circ \bar{\Phi}(z), \\ \bar{\Phi}(z) &:= \begin{pmatrix} z_1 \\ \vdots \\ z_{k-2} \\ \bar{\phi}_1(z_{k-1}, z_k; z_1, \dots, z_{k-2}) \\ \bar{\phi}_2(z_{k-1}, z_k; z_1, \dots, z_{k-2}) \end{pmatrix}, \end{aligned}$$

where $\bar{\phi} = (\bar{\phi}_1, \bar{\phi}_2)$ and its arguments $(z_{k-1}, z_k; z_1, \dots, z_{k-2})$ explicitly describe its dependency on the variables z_1, \dots, z_{k-2} . It can be observed that the properties (41) and (42) hold on the coordinate z obtained here.

Finally, the cases (i), (ii), and (iii) prove the theorem by induction. \square

REFERENCES

- [1] A. C. ANTOLIAS, *Approximation of Large-Scale Dynamical Systems*, Adv. Des. Control 6, SIAM, Philadelphia, 2005.
- [2] A. ASTOLFI, *Model reduction by moment matching for linear and nonlinear systems*, in Analysis and Design of Nonlinear Control Systems, A. Astolfi and L. Marconi, eds., Springer-Verlag, New York, 2008, pp. 429–444.
- [3] K. FUJIMOTO, *What are singular values of nonlinear operators?*, in Proceedings of the 43rd IEEE Conference on Decision and Control, IEEE Press, Piscataway, NJ, 2004, pp. 1623–1628.
- [4] K. FUJIMOTO AND J. M. A. SCHERPEN, *Eigenstructure of nonlinear Hankel operators*, in Nonlinear Control in the Year 2000, A. Isidori, F. Lamnabhi-Lagarigue, and W. Respondek, eds., Lecture Notes in Control and Inform. Sci. 258, Springer-Verlag, Paris, 2000, pp. 385–398.
- [5] K. FUJIMOTO AND J. M. A. SCHERPEN, *Balancing and model reduction for nonlinear systems based on the differential eigenstructure of Hankel operators*, in Proceedings of the 40th IEEE Conference on Decision and Control, IEEE Press, Piscataway, NJ, 2001, pp. 3252–3257.
- [6] K. FUJIMOTO AND J. M. A. SCHERPEN, *Nonlinear input-normal realizations based on the differential eigenstructure of Hankel operators*, IEEE Trans. Automat. Control, 50 (2005), pp. 2–18.
- [7] K. FUJIMOTO AND J. M. A. SCHERPEN, *Singular value analysis and balanced realizations for nonlinear systems*, in Model Order Reduction: Theory, Research Aspects and Applications, W. H. A. Schilders, H. A. van der Vorst, and J. Rommes, eds., Springer-Verlag, New York, 2008, pp. 251–272.
- [8] K. FUJIMOTO AND D. TSUBAKINO, *Computation of nonlinear balanced realization and model reduction based on Taylor series expansion*, Systems Control Lett., 57 (2008), pp. 283–289.
- [9] K. GLOVER, *All optimal Hankel-norm approximations of linear multivariable systems and their l^∞ -error bounds*, Int. J. Control, 39 (1984), pp. 1115–1193.
- [10] W. GRAY AND E. VERRIEST, *Balancing near stable invariant manifolds*, Automatica, 42 (2006), pp. 654–659.
- [11] W. S. GRAY AND J. M. A. SCHERPEN, *State dependent matrices in quadratic forms*, Systems Control Lett., 44 (2001), pp. 219–232.
- [12] J. HAHN AND T. F. EDGAR, *Reduction of nonlinear models using balancing of empirical gramians and Galerkin projections*, in Proceedings of the American Control Conference, IEEE Press, Piscataway, NJ, 2000, pp. 2864–2868.
- [13] P. HEUBERGER, *A family of reduced order models, based on open-loop balancing*, Ident. Model. and Contr., 1 (1990), pp. 1–10.
- [14] A. J. KRENER, *Reduced order modeling of nonlinear control systems*, in Analysis and Design of Nonlinear Control Systems, A. Astolfi and L. Marconi, eds., Springer-Verlag, 2008, pp. 41–62.
- [15] H. N. K. V. FERNANDO, *Singular perturbational model reduction of balanced systems*, IEEE Trans. Automat. Control, AC-27 (1982), pp. 466–468.
- [16] S. LALL, J. E. MARSDEN, AND S. GLAVASKI, *A subspace approach to balanced truncation for model reduction of nonlinear control systems*, Int. J. Robust Nonlinear Control, 12 (2002), pp. 519–535.
- [17] R. LOPEZLENA AND J. M. A. SCHERPEN, *Energy functions for dissipativity-based balancing of discrete-time nonlinear systems*, Math. Control Signals Systems, 18 (2006), pp. 345–368.
- [18] J. W. MILNOR, *Morse Theory*, Ann. Math. Stud. 51, Princeton University Press, Princeton, NJ, 1963.
- [19] J. W. MILNOR, *Topology from the Differential Viewpoint*, Princeton University Press, Princeton, NJ, 1965.
- [20] B. C. MOORE, *Principal component analysis in linear systems: Controllability, observability and model reduction*, IEEE Trans. Automat. Control, AC-26 (1981), pp. 17–32.

- [21] A. J. NEWMAN AND P. S. KRISHNAPRASAD, *Computation for nonlinear balancing*, in Proceedings of the 37th IEEE Conference on Decision and Control, IEEE Press, Piscataway, NJ, 1998, pp. 4103–4104.
- [22] L. PERNEBO AND L. M. SILVERMAN, *Model reduction via balanced state space representations*, IEEE Trans. Automat. Control, AC-27 (1982), pp. 382–387.
- [23] M. REWIEŃSKI, *A Trajectory Piecewise Linear Approach to Model Order Reduction of Nonlinear Dynamical Systems*, doctoral dissertation, MIT, Cambridge, MA, 2003.
- [24] J. M. A. SCHERPEN, *Balancing for nonlinear systems*, Systems Control Lett., 21 (1993), pp. 143–153.
- [25] J. M. A. SCHERPEN AND W. S. GRAY, *Minimality and local state decompositions of a nonlinear state space realization using energy functions*, IEEE Trans. Automat. Control, AC-45 (2000), pp. 2079–2086.
- [26] J. M. A. SCHERPEN AND W. S. GRAY, *Nonlinear Hilbert adjoints: Properties and applications to Hankel singular value analysis*, Nonlinear Anal. Ser. A Theory Methods, 51 (2002), pp. 883–901.
- [27] E. I. VERRIEST, *Time varying balanced realizations and nonlinear balancing*, in Model Order Reduction: Theory, Research Aspects and Applications, W. H. A. Schilders, H. A. van der Vorst, and J. Rommes, eds., Springer-Verlag, New York, 2008, pp. 213–250.
- [28] E. I. VERRIEST AND W. S. GRAY, *Flow balancing nonlinear systems*, in Proceedings of the 15th Symposium on Mathematical Theory of Networks and Systems (MTNS 2000), Perpignan, France, 2000, CD-ROM.
- [29] E. I. VERRIEST AND W. S. GRAY, *Discrete-time nonlinear balancing*, in Proceedings of the 5th IFAC Symposium on Nonlinear Control Systems (St. Petersburg, Russia), Elsevier, Amsterdam, 2001, pp. 515–520.
- [30] E. I. VERRIEST AND W. S. GRAY, *Nonlinear balanced realizations*, in Proceedings of the 43rd IEEE Conference on Decision and Control, IEEE Press, Piscataway, NJ, 2004, pp. 1164–1169.
- [31] E. I. VERRIEST AND W. S. GRAY, *Geometry and topology of the state space via balancing*, in Proceedings of the 17th Symposium on Mathematical Theory of Networks and Systems (MTNS 2006), Kyoto, Japan, 2006, pp. 840–848, CD-ROM.
- [32] K. ZHOU, J. C. DOYLE, AND K. GLOVER, *Robust and Optimal Control*, Prentice-Hall, Upper Saddle River, NJ, 1996.